PHD

An Examination of corporate structure and performance

Tian, Hui

Award date:
2018

Awarding institution:
University of Bath

AN EXAMINATION OF CORPORATE STRUCTURE AND PERFORMANCE

HUI TIAN

Thesis submitted to the University of Bath

For the degree of Doctor of Philosophy

October 2018

Abstract

This research project is based on three individual essays which are inter-linked. Our research starts from the development of a firm's profitability forecasting model. Given the much shorter history of quantile regression and its infrequent use in finance and accounting compared to the popular ordinary-least-squares (OLS) regression, it is not clear whether stock prices have fully reflected the incremental information from quantile-regression profitability forecasts over and above the information contained in their OLS counterparts. This thesis examines the issue using a forecasting and a hedge portfolio analysis. We construct quantile-regression forecasts on an economy-wide and an industry-specific basis and compare them to their OLS counterparts out-of-sample. We verify that the quantile-regression forecasts are more accurate than their OLS counterparts on either basis. We further show that a hedge portfolio formed by contrasting the quantile-regression forecasts to their OLS counterparts, whether on an economy-wide or industry-specific basis, can earn an abnormal return. Our results hold for a number of new and traditional profitability measures and are not sensitive to various methodological and sample choices.

Following similar interests in the profitability forecasting model, we focus on loss-making firms in the second essay. We examine the effects of diversification on loss persistence. Diversified firms have higher probability of loss reversal than focused firms. Using various measures of the abandonment option, we find that diversified firms can liquidate their loss-making assets or segments more efficiently than focused firms so as to achieve profits in the following year. Additional tests suggest that the efficiency of the abandonment option for diversified firms can be dampened by the agency problem of over-investment. Our findings are robust to various agency problem proxies and our analyses are controlled for endogeneity.

In our third essay, we investigate the effect of firm structure on dividend payout ratio. Consistent with the substitute theory, we find that diversified firms have significantly higher payout ratio than focused firms. We develop our analysis based on two hypotheses which are agency problem hypothesis and efficient internal capital market hypothesis. Under the agency problem hypothesis, we find that diversified firms have much higher dividend payout ratio than focused firms when they are under the agency problem but no difference in payout ratio when they are out of agency problem. Under the efficient internal capital market hypothesis, we find that diversification increases the dividend payout ratio for firms that are financially constrained but not for financially unconstrained firms. These findings further confirm the substitute theory that dividend payment is used as a signal for firms with agency problems and financial constraints. Our results are robust under various additional tests of endogeneity problems.

Acknowledgement

I would like to express my greatest gratitude to Professor David Newton, University of Bath for being an extraordinary supervisor and a mentor. His profound knowledge, encouragement and constructive feedback throughout my research not only sharpened my academic research skills but also improved my personal development. Without your support, inspiration and guidance, I would not have been able finish this thesis.

For those many academic researchers who have helped me in inspiring my thoughts and providing me with invaluable suggestions during my research period, I would like to express my appreciation to all of you. My sincere and special gratitude goes to Dr. Andrew Yim from Cass Business School, City University of London for being a supportive and invaluable mentor and co-author since my master's study. I would like to sincerely thank Wini Huang and Kevin Amess of being my internal and external examiners. I really appreciate for your time spent on my PhD viva, preparing the reports and very useful suggestions to improve my research. By working on the comments, the paper has gained considerably.

I also would like to thank my dear colleagues at the University of Bath School of Management: Yulin Wu (Harry), and Qi Hu (Sherry). Likewise, I express my gratitude to my colleagues at the University of Nottingham: Liang Liu, Haozhe Su, Zichen Yu (Jeff), and Kai Yao. I am grateful for your companionship, support, joy and unconditional friendship.

Finally, my greatest gratitude goes to my family - Mother and Father for their love and support. Without their support and encouragement, I could never have completed my research.

Table of content

List of Tables

List of Appendices

List of Figures

Abbreviations

AB_D	Abandonment Option Activities Dummy Variables
BV	Common/Ordinary shareholder's equity
CAPX	Capital Expenditure
CbOP_BS	Cash-based Operating profitability (Balance-sheet Approach)
CbOP_CF	Cash-based Operating Profitability (Cash-Flow Approach)
CFO	Cash flow from operating activities
CHE	Cash Holding
COVERAGE	Coverage Ratio
CRD	Credit Rating
DIFF	The difference of predicted profitabilities based on quantile regression over OLS regression
DIV	Dividend Payout Ratio
DIVD	Total Dividend
DIVDUM	Dividend paying dummy variable
DIVERSITY_D	Diversity dummy variable
EMP	Number of Employees
FCF	Free Cash Flow
FF12	Fama and French 12 Industry Classifications
FIRST_LOSS	First loss dummy variable
GDP	Gross Domestic Product
GP	Gross Profitability
GSL	Growth in sales

HERF(SALE)	Herfindahl Index by Sales
HERF(TA)	Herfindahl Index by Total Assets
LEVERAGE	Leverage level
LOSS_SEQ	Loss Sequence
MV	Market value of firm
NI	Income Before Extraordinary Items Available for Common Equity
NOA	Net operating assets
OP	Operating Profitability
OPINC	Operating Income After Depreciation
PAST_ROA	Average Return on Assets in the past five years
PAST_CFO	Average Cash Flow from Operating Activities in the past five years
PAST_ACCR	Average Accruals in the past five years
PNDIV	The fraction of all firms in the industry that are diversified firms
PSDIV	The fraction industry sales accounted for by diversified firms
RD	Research and Development Expenses
RE	Retained Earnings
REVERSAL	Loss Reversal
RNOA	Return on net operating assets
ROA	Return on asset
ROE	Return on equity
SALES	Sales
SIC	Standard Industry Classification
SIC2	2-digit Standard Industry Classification

SIZE	Firm Size
SPI	Special Items
T_Q	Tobin's Q
TA	Total Assets

Chapter 1

Main introduction

1.1 Background and research motivation

How much do investors and analysts know about firm performance and behaviour based on their fundamental characteristics? This might be the most interesting and widely-explored topic that both academic researchers and practitioners have dedicated their time to. This thesis explores this interesting topic by choosing three different but inter-related perspectives on firm performance. In the first essay, we advocate the superiority of quantile regression (least absolute deviate) over OLS regression (least squares) in firm profitability forecasting accuracy. In the second essay, we focus on the performance of loss-making firms. Referring to real option theory, we explain the difference in firms' loss reversal abilities under different structures, diversified firms vs. focused firms. In the third essay, we consider how firms distribute their profits, relating the 'dividend puzzle' to empirical evidence from firms with different structures.

Motivation for Essay 1

Investors and analysts seek the best predictions of firms' performances in order to decide on their most appropriate investment strategies. Firms' profitabilities are discussed in the literature in relation to mean reversion (Lev 1969, Brooks and Buckmaster 1976, Freeman, Olson and Penman 1982, Collins and Kothari 1989, Ou and Penman 1989, Easton and Zmijewski 1989, Penman 1991, Elgers and Lo 1994, Basu 1997, Fama and French 1995, Fama and French 2000). In a competitive market, the advantages of firms with extraordinarily high levels of profitability will, in time, be diluted by their competitors. Similarly, firms with depressed or negative earnings will make efforts to improve or leave the market in order to avoid losses and failure. As a result, the rate of return on investment converges in all industries under competition for established products or services (Stigler 1963). Researchers and practitioners then have a benchmark based on past profitability to predict future profitability. To best make the prediction, subsequent studies explore methods of forming this in-sample benchmark estimation: including comparing an industry-specific model with an economy-wide model (Fairfield et al. 2009), aggregate level forecasting with segment level (Shroeder and Yim 2017) and a firm-specific time-series forecasting model (Lev 1969, Freeman, Ohlson, and Penman 1982). Most in-sample estimations in such studies are based on OLS regression. In statistics, central tendency refers to the statistical measure that identifies a single value represneting the whole distribution (Gravetter and Wallnau 2016). However, given the rich literature documenting that firms' profitabilities are negatively skewed (Basu 1997, Givoly and Hayn 2000, Konstantinidi and Pope 2016), mean estimation via OLS regression may produce an inappropriate estimation of the central tendency of the profitability sample.

Quantile regression, an alternative approach based on the least absolute deviation method, was developed four decades ago by Koenker and Bassett (1978). This method is widely used in application areas such as medicine, survival analysis, and economics (Yu et al. 2003). Its applications in finance and accounting include return forecasting, portfolio analysis, risk measurement, and forecasting risk in earnings (Pohlman and Ma 2010, Basset Jr and Checn 2001, Lauridsen 2000, Konstatinidi and Pope 2016). Unlike OLS regression, it is not sensitive to outliers (Chen et al. 2008) and, particularly, median regression as its special case has found a role in several areas (Yu et al. 2003). Despite the apparent advantage of quantile regression, applications in finance and accounting remain uncommon. Owing perhaps to insufficient understanding by market participants, it is not clear whether stock prices have fully reflected the incremental information from quantile regression profitability forecasting over OLS. In Essay 1, we conduct a comprehensive comparison of median forecasting by quantile regression and mean forecasting by OLS regression, answering the following research questions:

Research questions for Essay 1:

1. Under the concept of mean reversion, how well does median estimation using quantile regression perform, compared to mean estimation quantile regression, in forecasting a firm's profitability?

2. Is the above conclusion sensitive to the profitability measures used?

3. How well does quantile regression outperform OLS regression, long term forecasting vs. short term?

4. Is the above sensitive to the forecasting model used?

5. Using simulation, how superior is quantile regression over OLS regression when related to the characteristics of the sample, such as skewness and kurtosis?

6. Is mean reversion at industry level or economy-wide level? Is an industry-specific model more accurate than an economy-wide model?

7. Is analyst forecast bias related to the distribution of the profitability? In other words, are analysts aware of the incremental advantage of quantile regression over OLS in forecasting?

8. Do investors consider the superiority of quantile regression over OLS in pricing stocks?

In order to address these questions, Essay 1 performs a series of empirical analyses. We build on various forecasting models to compare the forecasting accuracy between quantile regression and OLS regression, using actual firms' profitabilities. Robustness analyses provide a comprehensive picture of the usefulness of quantile regression in forecasting. We present evidence of the incremental information of quantile regression in analysts' forecasting bias analysis and hedge portfolio tests.

Motivation for Essay 2

In essay 2, we are interested in loss-making firms. In the mean reversion literature, negative changes in earnings and extreme changes reverse faster compared to changes with profitable firms (Elgers and Lo 1994, Fama and French 2000). This difference in earnings persistence between loss-making and profitable firms has been discussed and explored in subsequent studies. Hayn (1995) tests the informative of positive and negative earnings to the stock returns. She finds that negative earnings have much lower explanatory power for stock returns than positive earnings. She proposes using real option theory to explain this phenomenon, whereby loss-making firms have the option of abandonment to prevent persistent losses. Pinnuck and Lillis (2007) find that reporting accounting losses act as "an heuristic trigger for firms to exercise the abandonment option" and drop unproductive investments. This study provides solid evidence of the abandonment option applied by loss-making firms rather than the theoretical arguments in prior studies. A recent study by Lawrence et al. (2017) provides direct evidence that curtailments are important factors leading to the low persistence of losses. Inspired by these studies, my second essay pushes the idea further by investigating the efficiency of the exercising abandonment options - stated as an unsolved problem by Pinnuck and Lillis (2007).

We are interested in the efficiency (or otherwise) of exercising that abandonment option as related to firm diversification. The finance literature documents the importance of firm diversification in exercising options. Bernardo and Chowdhry (2002) emphasise the role of growth opportunities in accounting for the diversification discount. They argue that focused firms have more options to expand while diversified firms may have already exhausted these. In opposition to this, we argue that diversified firms have a selection of options (segments or business lines) making it easier and more efficient to liquidate their loss-making assets while focused firms only perform in the single industry making it harder for them to liquidate without closing the whole firm.

Research questions for Essay 2:

1. Are there any differences in loss reversal probabilities between diversified firms and focused firms when they are suffering losses?

2. How should we define abandonment options among loss-making firms?

3. Do diversified firms exercise their abandonment options more easily (frequently) than focused firms?

4. Do diversified firms exercise their abandonment options more efficiently?

5. Does the agency problem insert obstacles when firms try to exercise their options?

Building on the classic loss reversal profitability model of Joos and Plesko (2005), Essay 2 provides direct evidence of the explanatory power of firm diversification on the efficiency of exercising abandonment options.

Motivation for Essay 3

In essay three, we are interested in exploring differences in how firms distribute their earnings: specifically, dividend payouts related to firm diversification. There are common literatures shared between firm diversification and dividend policy. Diversified firms may be valued at a discount due to the agency problems (Berger et al., 1995) with diversified firms suffering higher agency problems due to conflicts between managers and shareholders. Although dividend irrelevance theory has been proposed following Miller and Modigliani (1961), various studies attempt to solve this puzzle using different theories: signalling, tax issues, and the catering hypothesis. Some studies argue that agency theory is one of the main issues affecting a firm's dividend policy

(Easterbrook 1984, La Porta et al. 2000). Dividend payout is used as a tool to distribute excess free cash flows to avoid their discretionary use by managers. Therefore, in Essay 3 we are interested in how firm diversification affects dividend policy.

Research questions for Essay 3:

1. Do diversified firms have different dividend payouts compared to focused firms?

2. If they do, is this related to agency problems?

3. How does this difference in dividend payout relate to internal capital?

Therefore, in the third essay, first we propose a regression test to show if there are any differences in dividend payouts between diversified firms and focused firms. Then we further test the agency problem hypothesis and internal efficient capital market hypothesis to provide more evidence of the relationship between dividend payout and firm diversification.

1.2. Structure of dissertation

Chapter 2 reviews the important literature related to the three essays. Since these essays have independent topics, we keep the general and comprehensive background literature in Chapter 2 and put essay-specific literature in the essays. Thus, Chapter2 reviews the literature of mean reversion of profitability (background literature for Essay 1), firm diversification (background literature for Essay 2&3), and dividend policy (background literature for essay 3).

Chapters 3, 4, and 5 are my three main pieces. Chapter 3 investigates model forecasting accuracy comparison between quantile regression and conventional OLS regression. Analyses here are model forecasting accuracy comparisons using various profitability measures, and hedged portfolio tests.

Chapter 4 examines the performance of loss-making firms by firm diversification characteristics. Based on the loss reversal model, we compare the loss reversal probability between diversified firms and focused firms and use real option theory to interpret my findings. We also group my sample by various agency problems to present further evidence.

Chapter 5 explains the famous dividend puzzle by investigating dividend payout with firm diversification. Agency problems and efficient internal market hypotheses are tested in the analyses. Like Chapter 4, Chapter 5 conducts various econometric techniques to eliminate endogeneity problems.

Chapter 6 draws conclusions from the three essays.

Chapter 2

Literature review

2.1 Mean reversion and profitability forecasting models

2.1.1. Quantile regression

Quantile regression has long been considered an attractive method in application areas such as medicine, survival analysis, and economics (Yu et al. 2003). Its applications in finance include return forecasting, portfolio analysis, and risk measurement (Pohlman and Ma 2010; Bassett Jr and Chen 2001; Lauridsen, 2000). Recent applications in accounting include forecasting risk in earnings (Konstantinidi and Pope 2016).

We propose constructing point forecasts of profitability using quantile regression, as opposed to the prevalent practice of using OLS regression. Specifically, we focus on the quantile regression for τ = 0.5 (i.e., the 50^{th} percentile), which is also referred to as median regression. This special case of quantile regression uses the absolute error loss criterion, as opposed to the squared error loss criterion upon which OLS regression is based. Median regression has the advantage of being more robust to outlier effects than OLS regression (Cameron and Trivedi 2005). Similarly, quantile regression is a more robust alternative for accommodating dependent variables with skewed distributions (Olsen et al. 2012). It is well-documented that firms' earnings are negatively skewed (Basu, 1997; Givoly and Hayn, 2000; Konstantinidi and Pope, 2016). This makes the mean estimation by OLS regression less appropriate for capturing the central tendency of the earnings distribution.

2.1.2. Profitability and mean reversion

Profitability measures, such as ROE and RNOA, are summary indicators of a firm's performance. Freeman et al. (1982) show that the mean in ROE can be expressed by a regression and establish that extreme ROEs are not as persistent as moderate ones. Fama and French (2000) provide evidence that mean reversion in firm profitability is a robust phenomenon and suggest that changes in profitability and earnings are to some degree predictable. In a simple partially adjusted model using US data, they find an estimated rate of mean reversion around 38% p.a. Similar results are documented by Allen and Salim (2005) who report a mean reversion rate of 25% p.a. in the UK market. We follow Fairfield et al. (2009) in using a forecasting model that captures the mean-reversion pattern of profitability conditional on the deviation of a firm's profitability from the median profitability benchmark (Fama and French 2000; Freeman et al. 1982).

Besides ROE and RNOA, we consider three alternative measures of profitability in our analysis. They are the gross profit, operating profit, and cash-based operating profit, deflated by the total assets lagged by one year. *Gross profit* is the sales minus the cost of goods sold. *Operating profit* is defined as the gross profit minus the selling, general, and administrative expenses reported (i.e., the Compustat-adjusted selling, general, and administrative expenses with the expenditures on research and development subtracted in order to undo this adjustment by Compustat). *Cash-based operating profit* is obtained by purging accruals from the operating profit. We consider two versions of cash-based operating profitability, depending on whether the balance-sheet or the cash-flow approach is used to convert operating profitability to a cash basis. Table 3.1 summarizes the definitions of the profitability measures examined in this study, which are consistent with prior studies (Novy-Marx, 2013; Ball et al. 2015, 2016; Fairfield et al. 2009).

Arguably, gross profitability, operating profitability, and cash-based operating profitability are cleaner measures of economic profitability than ROE and RNOA. Lower down the income statement, the net income for ROE and the operating profit after depreciation and amortization for RNOA are more polluted by financial reporting discretion, unlike gross and operating profits. Although the gross profit is cleanest in this sense, certain items farther down the income statement are not pure noise. Specifically, the reported selling, general, and administrative expenses, like the cost of goods sold, represent to a large extent expenses incurred to generate the current revenue. In contrast, the expenditures on research and development concern generating future revenues and are more discretionary in nature. Ball et al. (2015) show that the operating profitability, which has selling, general, and administrative expenses but not research and development expenditures subtracted off, can explain the cross section of stock returns even better than gross profitability. Ball et al. (2016) show that the cash-based operating profitability as an alternative free from accounting accruals adjustments can explain the cross section of stock returns even better than the operating profitability.

Being cleaner measures of a firm's economic profitability, the three measures are likely to have a closer tie to the industry membership of the firm. Therefore, they are especially relevant to our analysis comparing the industry-specific and economy-wide approaches to profitability forecasting.

2.1.3. Industry effects and broad industry classification

Early research documenting the mean reversion of profitability is often based on time series models fitted to individual firms separately (Brooks and Buckmaster 1976; Freeman et al. 1982; Lipe and Kormendi 1994; Penman 1991;). To minimize survivorship bias and increase the statistical power of tests, Fama and French (2000) use year-by-year cross-sectional regressions to establish evidence of mean-reverting profitability. None of these studies has considered the role of industry membership, as if it is irrelevant to firm profitability and its mean reversion pattern.

Consistent with this, Brown and Ball (1967) and Rumelt (1991) find only weak effects of industry difference in earnings. Moreover, Mueller and Raunig (1999) find that the profit rates of individual firms vary widely and do not converge to industry benchmarks even over a long period. Barber and Lyon (1996) also report that industry effects cannot effectively explain the future abnormal firm performance.

Other research, however, argues that industry membership is a fundamental determinant of firm profitability (Schmalensee 1985) and that industry differences affect the mean reversion of profitability (Gebhardt et al. 2001). Cheng (2005) finds that the increase in industry abnormal ROE is related to industry level factors such as industry concentration, industry level barriers to entry, and industry conservative accounting factors. Bou and Satorra (2007) observe significant and permanent differences in profitability between industries. Soliman (2004) show that industry-adjusted DuPont components of profitability is useful to predicting future profitability. McGahan and Porter (1997) also document the importance of industry effects on the variance of firm's profitability.

Fairfield et al. (2009) explore the potential of improving the accuracy of profitability forecasts using industry-specific OLS regressions. They find little forecast improvement over the economy-wide approach. In contrast, Schröder and Yim (2017) find industry effects in profitability forecasting for single-segment firms under various broad industry classifications. In comparing different forecasting approaches, we take advantage of their discovery of the importance of using a broad industry classification.

2.1.4 Profitability forecast versus current profitability

Gross profitability, operating profitability and cash-based operating profitability have received considerable attention in the literature because of their predictive power in explaining the cross section of stock returns (Novy-Marx 2013; Ball et al. 2015, 2016; Fama and French 2015, 2016, 2017; Akbas et al. 2017). This literature focuses on current profitability and its relation to stock return in the following year.

Current profitability may be viewed as the random-walk forecast of future profitability. Accordingly, the current profitability's relation to the stock return has a valuation interpretation. A higher current profitability is likely to result in a higher future profitability and a higher valuation of the stock, leading to a higher stock return.

Our interest in the profitability measures comes from their potential for valuation. Because valuation is forward-looking in nature, this study focuses on the forecasts of the measures, rather than their realized current levels.

2.2 Firm diversification

We review the diversification discount literature for Chapter 4 and Chapter 5.

2.2.1 Why diversify

Martin and Sayrak (2003) summarise three theoretical perspectives that explain why firms choose to diversify: a market-power view, a resource-view and an agency view. The market-power view serves the purpose of firm's profit maximization. Montgomery (1994) and Villalonga (2000) summarize three ways in which conglomerates can use to yield power: "deep-pocket" where a diversified firm can use its profits from one segment to support predatory pricing activities in another segment in a different market industry. Mutual forbearance, where competitors meeting in multiple markets recognize their interdependence and competition becomes less aggressively (Bernheim and Whinston 1990). Reciprocal buying, where smaller competitors are squeezed or eliminated by the firms engaging in reciprocal buying with other big firms.

The agency view focuses on the incentives of managers engaging in diversification strategies. First, managers can receive higher power and prestige (Jensen 1986, Stulz 1990) or higher managerial compensation (Jensen and Murphy 1990) by running larger firms. Amihud and Lev (1981) argue that a diversification strategy is not beneficial to shareholders by reducing the risk of the entity, since shareholders can achieve on their own the preferred degree of risk in their own portfolios. However, diversification can reduce the risk of the manager's undiversified personal portfolios ("employment risk" such as risk of losing jobs and personal reputations). Furthermore, managers can entrench themselves and reduce the probability of being replaced by shareholders through making manager-specific investments which require specific skills (Shleifer and Vishny 1989).

Early strategic management studies support diversification in a resource-based view. Diversification is an efficient form for organizing economic activities (Penrose 1959). These efficient economies of scale appear not only in the scope in production, distribution and marketing channels (Teece 1980, Teece 1982), but also from a managerial perspective such as utilizing financial and legal employees to support a different segment in different industries (Wernerfelt and Montgomery 1998, Bodnar et al., 1999) and financial synergies such as earnings smoothing. There is no clear evidence about the overall value of firm diversification.

2.2.2. Diversification destroys value

Are diversified firms worth less than their counterpart specialized firms in the same industry? Lang and Stulz (1994) find that industry-adjusted Tobin's Q of diversified firms is on average lower than single segment firms. Instead of Tobin's Q, a selection of the research documents evidence that diversified firms are priced at a discount of approximately 10%~ 15% relative to non-diversified companies in their industries by using the asset-based and the sales-based multiple methods (Berger and Ofek, 1995, Servaes and Lins 1999, and Lamont and Polk 2000). Berger and Ofeck (1995) apply this multiple valuation method to examine whether diversification creates or destroys firm value[1]. They measure the percentage difference between a firm's total value and the imputed value which is the sum of values for its segments as stand-alone firm. Later empirical studies apply this method and find the diversification discount phenomenon in their studies (Servaes, 1996; Stowe and Xing, 2006; Borghesi et al., 2007; Hoechle et al., 2012; Kuppuswamy and Villalonga, 2016).

Based on the finding for a diversification discount, studies focus on the potential causes of the poor diversification performance. The main potential reason is the effect of the internal capital allocation of diversified firms. Theoretical studies call the internal transfer of resources within diversified firms a cross-subsidization. There are two sides of cross-subsidization. It can be efficient if it helps the firm overcome financial constraint, while can be inefficient if it causes firms to misallocate their funds. Here, we focus on the inefficient part as it is related to the diversification discount literature.

A substantial body of empirical papers support the capital misallocation hypothesis. Evidence has been found that diversified firms tend to overinvest in business lines with poor investment opportunities (Berge and Ofek 1995, Stulz 1990, Scharfstein 1998). In particular, a study by Shin and Stulz (1998) concludes that the capital expenditure made by a diversified firm's segment not only depends on its own cash flows but also on the cash flows of other segments. In addition, the capital expenditures by the segments of firms with high diversity level are not as sensitive as to their cash flows than the capital expenditures of comparable focused firms. More importantly, the

[1] See Appendix for a detailed introduction of this method.

sensitivity of a segment's capital expenditure to the cash flow of other segments is not related to if the investment opportunities (Tobin's Q) are better than those of the firm's other individual segments. This may cause the issue of over-investment and inefficient investment. Meyer, Milgrom, and Robert (1992) note that poorly performing business segments cannot have a negative value if operated on their own (as the independent entities) but can if being cross-subsidized among diversified firms. Therefore, they predict that loss-making segment or business create more value loss in diversified firms than they would as stand-alone firms.

Some literature finds that the inefficiency of cross-subsidization is caused mainly by agency problems, since the agency problem is one of the main incentives for firms choosing to diversify. Jensen (1986) points out that managers are more likely to undertake negative NPV (net present value) projects by using "unused borrowing power" and excess free cash flows. Since business lines are more likely to access higher amounts of free cash flows as a part of diversified firms than on their own, diversified firms tend to invest more in negative NPV projects than their segments would if they were investing as independent entities. A recent study by Denis, Denis and Sarin (2012) finds that there is a negative relationship between the diversity level and managerial equity ownership. They find a correlation between decreases in diversification with financial distress, management turnover and external control threats. These findings suggest that agency problems are the main reasons causing firms' value-reducing diversification strategies. A similar study by Rajan, Servaes and Zingales (2000) shows that diversification mainly causes resources to flow to inefficient investments due to power struggles in firm's divisions. Scharfstein and Stein (2000) provide evidence of the subverting of the internal capital budgeting allocation by the rent-seeking segment managers.

2.2.3. Concerns in diversification discount

Early strategic papers point out that the decision to diversify is not random. There are systematic differences between the firms choosing to diversify and those choosing not to diversify (Lemelin, 1982; MacDonald, 1985; Montgomery and Hariharan, 1991; Merino and Rodrigues 1997; Silverman 1999). Lang and Stulz (1994) point out that firms that choose to diversify are poor performers relative to firms that choose not to. Several studies support this view by showing that diversified firms have already valued at a discount before they diversify (Hyland 1999; Campa and Kedia 2002; Villalonga 1999). Graham et al. (1999) document that an average diversification discount of approximately 15% is related to the acquired segments which were also discounted before acquisition as focused firms. Therefore, there is a need to eliminate the self- selection problem when evaluating treatment effects with non-experimental data.

Campa and Kedia (2002) and Villalonga (1999) apply various econometric techniques to reduce the issue of self-selection bias. Methods include propensity score matching, Heckman's two-stage model and instrumental variables. While applying these techniques, various firm level differences between diversified firms and focused firms are controlled for: size, industry growth rate, and variables deflated by sales such as capital expenditures, profitability and research and development expenses. After applying these methods to eliminate the self-selection problem, the diversification discount disappears or even becomes a premium.

There are other studies challenging the existence of diversification discount. First, instead of using the Compustat segment database, Villalonga (2003) find a diversification premium by using a new Census database covering the whole U.S economy at the establishment level. A possible explanation raised by Villalonga is that segment data measure purely unrelated diversification while established data measure both related and unrelated diversification. In the spirit of the early strategic papers addressing the related and unrelated diversification effect on firm performance

(Rumelt 1974), Villalonga’s finding can be interpreted as showing that there is a premium to related diversification but a discount to unrelated diversification (such as conglomerate).

Other studies, such as the one by Mansi and Reeb (2002), question the existence of a diversification discount by challenging the method of calculating excess values based on the sales and asset based multiple. Since this method define the market value of a firm as the sum of the market equity value and book value of equity, they find that such diversification discount disappears in all-equity firms. They argue that the use of book value of debt to calculate excess value generates a downward bias for diversified firms.

2.2.4. Diversification creates value

The benefits of diversification giving rise to a diversification premium is supported by different views in the literature. Corporate finance literature focuses on the efficient internal capital market view that the corporate headquarters can transfer and allocate valuable resources to competing projects in an internal capital market. Stein (1997) points out that this “winner-picking” funds shift within firm is efficient and value-enhancing when the overall firm is facing credit constraints (when not all the positive NPV projects can be financed). The notion of “winner-picking” is based on several early papers (Williamson 1975, Donaldson 1984) which argue that cash flows in diversified firms are exposed to internal competition rather than being allocated automatically. Similar studies show evidence for investment interdependence. Shin and Stulz (1996) indicate that within diversified firms the investments of small divisions are strongly associated with the cash flows of other divisions. The phenomenon of “intracompany liquidity spillovers” in the U.S. oil industry is demonstrated in Lamont (1997). The major oil companies suffering an oil price decline in 1986 cut investments across the board even in their non-oil-related divisions to deal with the decline in cash flows.

2.3. Dividend literature

Why do some firms pay dividends while others do not? This has been of interest since the dividend irrelevance paper by Miller and Modigliani (1961) was published. The M&M theorem suggests that in a perfect capital market, the dividend decision is irrelevant to the firm's value. The reason is that higher dividend payouts lead to lower retained earnings and capital gains with shareholder's wealth unchanged.

2.3.1. Dividend smoothing and dividend puzzle

So-called dividend smoothing is probably the most widely-documented phenomenon in early studies. Lintner (1956) conducted a small study on how U.S. managers make decisions on dividends, interviewing managers from 28 firms using 15 variables that might impact dividend decisions. He found that a change in a firm's earnings has the most impact on the size of dividend since firms tend to make partial adjustments in payout ratio each period to the direction of their target payout ratio rather than paying dividends with significant changes periodically. Managers tend to believe strongly that a premium is added by the market on firms with a stable and consistent dividend policy. Therefore, dividend smoothing is implemented by the managers to avoid a fluctuating dividend streams. Subsequent studies support Lintner's findings on managers' preference for stable dividend policy (Brittain 1964, Benartzi, Michaely, and Thaler 1997, Brav et al., 2005). Fama and Babiak (1968) show that changes in the dividends are closely associated with a firm's target payout, current or previous earnings, and previous dividend. Brav et al. (2005) find that managers are willing to avoid cutting the dividend at the cost of raising costly external funds or even foregoing positive NPV projects.

Lintner's (1956) dividend smoothing phenomenon and methodology inspired a series of empirical studies examining cross-sectional differences and using broader sample ranges.

Michaely and Roberts (2011) suggest that the propensity to smooth dividends is closely related to ownership structure among UK firms. Specifically, they find that private firms smooth dividends much less than counterpart public firms. This indicates that public capital markets play a key role in the propensity of firms to conduct dividend smoothing. Dewenter and Warther (1998) compare dividend policies between U.S. firms and Japanese firms. They find that Japanese firms (keiretsu members) smooth less than U.S. firms since they face less information asymmetry and agency conflicts. Aivazian, Booth, and Cleary (2006) find that firms with bond ratings follow the pattern of dividend smoothing in Lintner (1956), while firms with no bond ratings show much less dividend smoothing behaviour.

Although these stylized facts are well-documented in literature, it is still not understood how and why firms choose their dividend policies. The puzzle becomes deeper in those countries such as the U.S. where dividend receipts are taxed more heavily than capital gains. Therefore, we summarize a few hypotheses that could explain this dividend puzzle in different perspectives.

2.3.3. Signalling hypothesis

Like the dividend smoothing literature, the signalling effect of dividends is addressed in subsequent studies. The signalling hypothesis suggests that firms can communicate their views of future performance through paying dividends since managers have asymmetric information about the firms (Bhattachrya 1979, Miller and Rock 1985, John and Williams 1985, and Ambarish, John, and Williams 1987). There are two approaches to testing the relationship. One is to examine whether changes in dividends can predict changes in stock prices. A positive relation between dividend change and dividend announcement is documented in the literature (Aharony and Swary 1980, Asquith and Mullins 1983, Bajaj and Vijh, 1990, Kalay and Loewenstein 1985, Petit 1972). The other approach focuses on the relation between dividend changes and the changes in future accounting earnings (Watts 9173, Gonedes 1978, Penman 1983, DeAngelo et al. 1996, Benartzi et al. 1997, Nissim and Ziv 2001, Grullon, Michaely and Swaminathan 2002, Brav et al. 2005, Grullon et al. 2005, Denis and Osobov 2008, Braggion and Moore 2011). The signalling hypothesis assumes that the managers resist increasing (decreasing) dividends unless the increase (decrease) in earnings is persistent. Koch and Sun (2005) provide more direct evidence and find that investors do indeed interpret a change in dividend as a signal about the change of persistence of past earnings.

The signalling theory has been challenged by several scholars. DeAngelo, DeAngelo and Skinner (1996) use various model specifications but find no evidence that dividends signal superior future profitability. They argue that a dividend is not a reliable signal, through six possible explanations. (1). Current earnings are so informative about future earnings that they leave little additional content to dividends. (2). Dividends can be a free-cash-flow payout rather than signalling when managers cut off capital outlays. (3). Dividend increases are not favourable signals but instead are lagged responses to previous increases in earnings. (4). Managers send good dividend signals by mistake, while “these mistakes are understandable given the information available at the time”. (5). Managers tend to overestimate a firm’s future growth by sending the optimistic dividend signal. (6). Only modest cash commitments are made by managers when they increase dividends. This can undermine the reliability of the signal. Through a comprehensive analysis, their findings support (5) and (6). Specifically, (5) is consistent with the Jensen’s (1993) behavioural hypothesis. The theory suggests that managerial mindset and corporate culture normally lead managers to overestimate the firm’s growth perspective.

Other studies present different arguments against the dividend signalling hypothesis. Myers and Majluf (1984) suggest that the distribution of dividends can affect future earnings but not by signalling. Since firms prefer internal funds over external funds, dividend increase can reduce the amount of internal funds for investment. Consequently, firms may need to access costly external funds or forgo positive NPV projects, which can cause a decrease in future earnings. This relation can be also shown in the constant dividend growth model of Gordon (1962). With expected return of portfolio constant, high dividend payout should be offset by low expected earnings growth.

2.3.4. Dividend and agency hypothesis

Other studies suggest different motives behind dividend policy, converging with agency theory. Jensen's (1986) agency theory suggests that managers are more likely to undertake negative NPV projects by using unused excess free cash flows. Extensive studies show that dividends are used as a tool for distributing excess free cash flows in order to reduce agency costs. (Rozeff 1982, Easterbrook 1984, Lang and Litzenberger 1989, Grullon and Michaely 2004, DeAngelo, DeAngelo and Stulz 2006, Michaely and Roberts 2012). Thus, agency theory predicts that shareholders may prefer dividends over retained earnings.

In Esterbrook's work (1984), dividends play a monitoring role in mitigating the conflicts between managers and shareholders, while agency theory is not limited to conflicts between managers and shareholders but also applies to the inside shareholders (controlling shareholder) and the outside shareholders (minority shareholders). La Porta et al. (2000) introduce two competing models explaining how conflicts between corporate insiders and minority shareholders influence dividend policy. The outcome model states that dividend payout is the outcome of minority shareholders pursuing legal protection. Evidence is found that dividend payouts in civil-law countries are much lower than that in common-law countries. They claim that minority shareholders can use their legal power to force insiders to distribute excess cash. The other model is substitute model. Insiders pay dividends to establish a reputation for decent treatment of minority shareholders since insiders need to build a good reputation for the future fund-raising. These two models predict opposite impacts of the agency problem on dividend payouts: the outcome model predicts stronger minority shareholder rights should relate to higher payout while the substitute model predict the opposite.

2.3.5. Catering hypothesis

Yet other empirical studies focus on the role of investor demand in affecting dividend policy, which is known as the catering hypothesis. Firms paying dividends cater to a preference for dividends from heterogeneous clienteles (Baker and Wurgler 2004a & 2004b, Jayaraman and Sabherwal 2009, Li and Lie 2006). Here we have two interesting examples. Becker, Ivkovic and Weisbenner (2011) document retail investors preferring local stocks and older investors tending to hold dividend-paying stocks. They find that firms headquartered in areas where most populations are seniors tend to pay dividends, initiate dividends and have higher dividend yields. Desai and Jin (2011) find that “dividend-averse” institution shareholders are significantly less likely to hold shares in firms which pay high dividends.

2.3.6. Life-cycle theory

Life-cycle theory considers dividend payout patterns from a different perspective. This theory reflects a financial cycle and suggests that dividends are more likely to be paid by mature and established firms than by young and small firms. On the one hand, young firms may face financial constraints if they insist on paying dividends but simultaneously have abundant investment opportunities. On the other hand, mature and established firms have higher and stable profitability and fewer attractive investment opportunities and so are better placed to pay dividends. Evidence for life-cycle theory can be found in U.S. and developed markets (DeAngelo et al., 2006, Fama and French 2001, Grullon et al., 2002, Denis and Osbov 2008) as well as in emerging markets (Thanatawee 2011).

Chapter 3

Stock Returns and Profitability Forecasting by Quantile Regression

3.1. Introduction

In an efficient market, stock prices should fully reflect the prospects of firm profitability anticipated by the market, based on publicly available information. To formulate profitability forecasts as accurately as possible, sophisticated market participants are likely to resort to statistical methods. Ordinary-least-squares (OLS) regression is a very popular choice, if not the prevalent choice. The least squares method has a very long mathematical history dating back to 1795 (Courgeau 2012). Given its familiarity, educated investors should be able to figure out the implications of the information contained in OLS-regression forecasts of profitability.

In contrast, quantile regression, an alternative approach based on the least absolute deviation method, was developed only four decades ago by Koenker and Bassett (1978). Unlike the least squares method, the least absolute deviation method is not especially sensitive to outliers (Chen et al. 2008). In recent years, quantile regression, or median regression as its special case, has found a role in various areas (Yu et al. 2003). Despite the advantage of quantile regression, its applications in finance and accounting remain uncommon (see section 2.1 for more details). Owing to an insufficient understanding by market participants, it is not clear whether stock prices have fully reflected the incremental information from quantile-regression profitability forecasts over and above the information contained in their OLS counterparts.

This chapter examines the issue through two sets of analyses. In the forecasting analysis, we construct quantile-regression forecasts on an economy-wide and an industry-specific basis and compare them to their OLS counterparts out-of-sample. We verify that in terms of absolute forecast error, the quantile-regression forecasts are more accurate than their OLS counterparts on either basis. In this sense, quantile-regression forecasts are more informative than OLS-regression forecasts. Moreover, we find that industry-specific quantile-regression forecasts are most accurate among the four types of forecasts.

In the hedge portfolio analysis, we utilize the incremental information from the quantile-regression forecasts by sorting stocks according to the excess of the quantile-regression profitability forecasts over their OLS counterparts. Quantile-regression forecasts being more accurate than their OLS counterparts, investors should follow the investment guidance by the former rather than the latter. Therefore, we form a dollar-neutral hedge portfolio with a long position in the stocks of the large-excess firms and a short position in those of the small-excess firms. The incremental information from the quantile-regression forecasts is economically significant if investors can earn an abnormal return from the portfolio constructed with the information. We find significantly positive alphas based on the Carhart (1997) four-factor and the Fama and French (2015) five-factor asset pricing model. This suggests that the incremental information from the quantile-regression forecasts have not been fully impounded in stock prices, consistent with the hypothesis that quantile regression is much less familiar to investors than OLS regression.

To our knowledge, we are the first to examine whether quantile regression is more accurate than OLS regression in constructing point forecasts of profitability.[2] The motivation follows from the observation of the highly-skewed earnings distribution. In statistics, central tendency is described as the statistical measure that identifies a single value as representative of the whole distribution (Gravetter and Wallnau 2016. When a distribution is asymmetrical, the median is

[2] Despite the availability of methods to produce interval and density forecasts, point forecasts remain the most commonly used in practice. They are often easier to understand and act upon and are less costly to produce (Diebold 2015).

considered the more preferred measure of central tendency than the mean since the former is less affected by outliers and skewed data (Healey 2015). We investigate forecasting by quantile regression because it can produce median forecasts as opposed to mean forecasts produced by OLS regression.

We focus on three new profitability measures in our comparison of the different types of forecasts. They are the gross profitability (GP) defined by Novy-Marx (2013), operating profitability (OP) defined by Ball et al. (2015) and two versions of cash-based operating profitability (CbOP) defined by Ball et al. (2016). Novy-Marx (2013) finds that GP can explain most earnings-related anomalies. Ball et al. (2015), however, show that OP has a much stronger link with stock returns than GP. The usefulness of OP in explaining the cross section of stock returns has led to its inclusion as a new factor in the latest five-factor asset pricing model (Fama and French 2015; 2016; 2017). Adding to the success of OP, Ball et al. (2016) show that CbOP outperforms OP in predicting the cross section of stock returns, explaining two anomalies related to accruals and profitability measures that include accruals.

Besides the new profitability measures above, we also include return on equity (ROE) and return on net operating assets (RNOA) in our comparison. Prior research on profitability forecasting examines these conventional measures of profitability because they are the inputs to accounting-based valuation models (Fairfield et al. 2009; Schöder and Yim 2017). Their inclusion here facilitates the comparison of our results with prior research findings. It is also interesting to include ROE in its own right. This is the profitability measure used in the Hou et al. (2015) *q*-factor asset pricing model, whose performance is comparable to and sometimes even better than that of the Fama and French (1993) three-factor model and the Carhart (1997) four-factor model.

This chapter contributes to the broad literature on market anomalies and stock return predictability (e.g., Fama and French 2008; Richardson et al. 2010; Campbell 2015; Mclean and Pontiff 2016; Akbas et al. 2017; Green et al. 2017). We find evidence for quantile-regression profitability forecasts being more accurate than their popular OLS counterparts. Despite this, the low familiarity appears to have prevented the information uniquely contained in the quantile-regression forecasts from being impounded into stock prices. Consequently, a hedge portfolio formed by contrasting the quantile-regression forecasts to their OLS counterparts can yield an abnormal return.

We also contribute to the valuation literature by showing that profitability can be forecast more accurately by industry-specific quantile regression. Price multiples are widely used for firm valuation despite the availability of more sophisticated valuation methods (Roosenboom 2012; Imam et al. 2008; Asquith et al. 2005; Demirakos et al. 2004; Fernandez 2002). Liu et al. (2002) have examined the valuation performance of a long list of multiples. They find that multiples derived from forward earnings explain stock prices significantly well. In contrast, multiples based on historical earnings, cash flow, book value of equity, or sales do not perform as well. How useful forward earnings are to multiple-based valuation depends on the forecast accuracy. One could rely on financial analysts' professional skills to obtain high-quality earnings forecasts. However, analyst earnings forecasts are unavailable for many firms (Li and Mohanram, 2014). Besides, analysts also use forecasting models to assist in developing their forecasts. Therefore, it is important to know how model-based earnings forecasts can be constructed more accurately.[3]

[3] Forecasting earnings in practice is often equivalent to forecasting profitability (e.g., Li 2011; Chang et al. 2016). Data samples used to forecast earnings typically include firms of different sizes. Deflation is a technique to control for the size differences. Deflating an earnings measure by a certain size variable, such as book value of equity, net operating assets, or total assets, gives a profitability measure (Fairfield et al. 2009; Li et al. 2014; Schröder and Yim 2016).

Lastly, although we are not the first to compare the industry-specific and economy-wide approaches to profitability forecasting, we are first to compare them for quantile regression. Fairfield et al. (2009) investigate the two approaches and conclude that there is little improvement from using the industry-specific approach to forecast profitability. Schröder and Yim (2017) re-examine the topic utilizing segment reporting information. They find that the industry-specific approach produces more accurate forecasts for single-segment firms under broad industry classifications, such as the Fama-French 12-industry and the first-digit Standard Industry Classification (SIC). We take advantage of their discovery of the importance of using a broad industry classification. For the first time, we find forecast improvements for *all firms* when using industry-specific quantile regression, as opposed to economy-wide quantile regression, to construct profitability forecasts.

3.2. Research Design and Sample Selection

3.2.1. Forecasting analysis

Consistent with prior studies such as Fairfield et al. (2009) and Li and Mohanram (2014), we construct the profitability forecast for each firm-year in two steps. First, we estimate in-sample a forecasting model on a rolling basis using the data of all the firms for the previous ten years. For example, to forecast the profitability of a firm for year T, we first estimate the coefficients of a forecasting model using the data of all the firms from year $T-10$ to year $T-1$. Next, we apply the estimated coefficients from the in-sample regression to the current-year data of a firm to obtain the one-year-ahead profitability forecast of the firm.

We consider four forecasting approaches. The first approach uses the following forecasting model based on the *economy-wide OLS regression* specification studied in Fairfield et al. (2009):

$$x_{i,t} = \alpha_T + \beta_T x_{i,t-1} + \gamma_T D_{i,t} * x_{i,t-1} + \lambda_T PREDGSL_{i,t} + u_{i,t}, \tag{1}$$

where $t = T-10, \dots, T-1$. The dependent variable $x_{i,t}$, indexed by firm i and year t, stands for one of the profitability measures considered: GP, OP, CbOP, ROE, and RNOA. $D_{i,t}$ is a dummy variable equal to one if in year $t-1$, the profitability of firm i is below the median profitability of all firms and equal to zero otherwise. $PREDGSL_{i,t}$ is the predicted growth in sales, which is found to be useful for profitability forecasting. $u_{i,t}$ is the error term. The model parameters α_T, β_T, γ_T, and λ_T are indexed by year T to highlight that they are estimated for each year T using the previous ten years of data.

To construct *PREDGSL*, we use the following simple first-order autoregressive model estimated by OLS regression on an industry-specific basis:

$$g_{i,t} = \mu_{j,T} + \upsilon_{j,T} g_{i,t-1} + \epsilon_{i,t}, \tag{2}$$

where $g_{i,t}$ is the growth in sales of firm i in year t, $\epsilon_{i,t}$ is the error term, and $t = T-10, \dots, T-1$. The model parameters $\mu_{j,T}$ and $\upsilon_{j,T}$ are indexed by industry j and year T to highlight that the estimation is done on an industry-specific basis and for each year T using the previous ten years of data. The $PREDGSL_{i,T}$ for each firm-year (i,T) is set to the predicted value $m_{j,T} + n_{j,T} g_{i,T-1}$, where $m_{j,T}$ and $n_{j,T}$ are the estimated coefficients of the model parameters $\mu_{j,T}$ and $\upsilon_{j,T}$. We construct *PREDGSL* by OLS regression on an industry-specific basis since Fairfield et al. (2009) find that sales growth forecasts are more accurate when constructed this way, rather than on an economy-wide basis.[4]

[4] We verify that this also holds for our sample. We discuss in section 5 the robustness of our results to alternative ways to construct *PREDGSL*.

Our second forecasting approach, *economy-wide quantile regression*, uses the same model as specified in equation 1 except that the parameters (α_T, β_T, γ_T, λ_T) are estimated by quantile regression for τ = 0.5 (i.e., by median regression). In general, quantile regression estimates are obtained by minimizing the loss function $\rho_\tau(u)$ on the error term u (see the illustration in Figure 3.1). For τ = 0.5, the loss function becomes symmetric and equals $|u|$. The quantile regression estimates for this case are conditional median estimates. In our context, the estimated coefficients are given by

$$\underset{(\alpha_T,\, \beta_T,\, \gamma_T,\, \lambda_T)}{\operatorname{argmin}} \sum_{i,t} \left| x_{i,t} - (\alpha_T + \beta_T x_{i,t-1} + \gamma_T D_{i,t} * x_{i,t-1} + \lambda_T PREDGSL_{i,t}) \right|. \tag{3}$$

Our third forecasting approach, *industry-specific quantile regression*, uses the following industry-specific forecasting model:

$$x_{i,t} = \alpha_{j,T} + \beta_{j,T} x_{i,t-1} + \gamma_{j,T} D_{i,t} * x_{i,t-1} + \lambda_{j,T} PREDGSL_{i,t} + \varepsilon_{i,t}, \tag{4}$$

where $t = T-10, \ldots, T-1$, with the parameters $(\alpha_{j,T}, \beta_{j,T}, \gamma_{j,T}, \lambda_{j,T})$ estimated by quantile regression for τ = 0.5 on an industry-specific basis. For this forecasting model as well as the construction of $PREDGSL_{i,t}$ throughout the different forecasting approaches, the industries are classified based on the first-digit SIC. Schröder and Yim (2017) find that a broad industry classification like this better balances the bias from model misspecification and the sample size for industry-specific estimation, increasing the chance of finding a forecasting improvement. We examine whether this finding for OLS regression holds for quantile regression as well.

We compare industry-specific quantile regression to the fourth forecasting approach, *industry-specific OLS regression*, which estimates equation 4 by OLS regression on an industry-specific basis. This last comparison verifies whether quantile regression is more accurate than OLS regression not only on an economy-wide basis but also on an industry-specific basis.

Following prior research such as Li et al. (2014) and Fairfield et al. (2009), we use absolute forecast error (AFE) to measure the accuracy of a forecasting approach. Specifically, the AFE of forecasting approach A for a firm-year (i,T) is defined as the absolute difference between the actual profitability $x_{i,T}$ and the profitability forecast $E_A[x_{i,T}]$ constructed with forecasting approach A:

$$AFE_A(i,T) = |\, x_{i,T} - E_A[x_{i,T}] \,|. \tag{5}$$

For example, the profitability forecast constructed with the first approach (i.e., economy-wide OLS) is

$$E_{ew\_OLS}[x_{i,T}] = a_T + b_T x_{i,T-1} + c_T D_{i,T} * x_{i,T-1} + l_T PREDGSL_{i,T}, \tag{6}$$

where (a_T, b_T, c_T, l_T) are the economy-wide OLS estimates of the model parameters (α_T, β_T, γ_T, λ_T). Because the actual profitability is not part of the data used to construct the profitability forecast, the assessment by the AFE is said to be out-of-sample.

To assess the relative accuracy of two forecasting approaches, say, A and B, we compute the forecast improvement (FI) of approach A over B for a firm-year (i,T). This is defined as the difference in the AFE between the forecasts from the two approaches:

$$FI_{A,B}(i,T) = AFE_B(i,T) - AFE_A(i,T). \tag{7}$$

The FI would be positive if approach A has a lower AFE than approach B. To conclude on the relative accuracy of the two forecasting approaches, we perform tests on the mean as well as the median FI over all firm-years. Consistent with the framework of comparing predictive accuracy in Diebold and Mariano (1995), the test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test.

To ensure an equal-footing comparison, the same sample is used to construct the forecasts by the industry-specific and the economy-wide quantile regression approach. We first determine the sample used for the industry-specific quantile regression approach. This is the more restricted sample because whenever estimation is done on an industry-specific basis, we include only the industries with at least 100 firm-year observations in the rolling sample period of the previous ten years to avoid unreliable estimation[5].

3.2.2. Hedge portfolio analysis

To see whether investors can benefit from quantile-regression profitability forecasts, we perform a hedge portfolio analysis similar to Li et al's (2014) to link forecast accuracy to stock return predictability. The hedge portfolio in the analysis is constructed based on the relative accuracy of an improved approach versus a benchmark approach. Suppose that knowing the forecasts from the improved approach adds nothing beyond knowing the forecasts from the benchmark approach. It would not be possible to earn an abnormal return from a hedge portfolio constructed using the forecasts of the two approaches. However, if an abnormal return can be found for such a portfolio, the forecasts from the improved approach must contain economically important information not already reflected in the forecasts from the benchmark approach.

We construct the hedge portfolio for year T by sorting the sample of firms into quartiles based on

$$DIFF_{\mathrm{A,B}}\,(i, T+1) \; = E_{\mathrm{A}}\left[x_{i,T+1}\right] - E_{\mathrm{B}}\left[x_{i,T+1}\right]. \tag{8}$$

This measure captures the extent to which the profitability forecast of an improved approach A is higher than that of a benchmark approach B. Because approach A on average is more accurate than approach B, one should trust the guidance by the profitability forecast $E_{\mathrm{A}}\left[x_{i,T}\right]$ more than that by

[5] This is consistent with the Fairfield et al., (2009, p 158) by requiring at least 100 observations to generate a reliable estimation of the coefficients for both industry-specific and economy-wide forecasting models. We also tried different minimum number of observations in the in-sample estimation, the results are very close.

$E_{\mathrm{B}}\left[x_{i,T}\right]$. To take advantage of the better guidance, a dollar-neutral hedge portfolio can be constructed by shorting the lowest-quartile firms and going long the highest-quartile firms. Considering further that the implication of $DIFF_{\mathrm{A,B}}\,(i, T+1)$ is less likely to be reliable for those firms with $FI_{\mathrm{A,B}}(i,T) \leq 0$, we confine to the subsample of firms with positive FIs in sorting the firms into quartiles for the hedge portfolio construction. To avoid a look-ahead bias, the portfolio is formed each year at the end of June based on forecasts constructed with data available by the end of December of the preceding year. The portfolio is held for the following year to track the monthly returns.

We report the value-weighted raw and risk-adjusted returns of the hedge portfolio. The risk-adjusted return is the alpha from the time-series regression of the hedge-portfolio returns (*HEDGE*) onto the returns to a market portfolio of stocks (*MKT*) and factor-mimicking portfolios based on two asset pricing models. They are (i) the Carhart (1997) four -factor model:

$$HEDGE_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \epsilon_t, \tag{9}$$

where *SMB*, *HML*, and *UMD* denote the returns to the size, book-to-market, and momentum factors, respectively; (ii) the Fama and French (2015) five-factor model:

$$HEDGE_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CMA_t + \beta_5 RMW_t + \epsilon_t, \tag{10}$$

where *CMA* and *RMW* denote the returns to the investment and profitability factors, respectively.

3.3. Sample selection

Profitability forecasts for the forecasting analysis are constructed for the period from 1989 to 2016 because some measures require data from the cash flow statements available only from 1988 onwards. The forecasts used for the hedge portfolio analysis are from 1989 to 2014 since we need to track the returns of the hedge portfolios formed at the end of each following June and held for the next twelve months. We use the previous ten years of data to construct the profitability forecasts for a year. As the *PREDGSL* variable in the forecasting models requires ten earlier years of data to construct, the profitability forecasts for 1989 are constructed with data as far back as 1969.

We obtain accounting data of US firms from the Compustat North America annual fundamentals file and stock data from the Centre for Research in Security Prices (CRSP) monthly stock file on Wharton Research Data Services (WRDS). Only observations with non-missing SIC codes are retained. We exclude financial and utility firms (SIC from 6000 to 6799, or from 4900 to 4949) because they are highly regulated. In addition, the U.S. postal service (SIC 4311) and public administration (SIC 9000 or above) are also excluded.

To mitigate the effect of small denominators on the profitability and growth in sales measures, observations with total assets, net operating assets, or sales below USD 10 million or book value of equity below USD 1 million are excluded. To further mitigate the effect of mergers and acquisitions on the relation between current-year and lagged variables, we remove observations with growth in total assets, net operating assets, sales, or book value of equity exceeding 100%.

To reduce the influence of outliers, we remove observations with any of the profitability and sales growth measures exceeding 1 in absolute value. For the in-sample estimation of the forecasting models, we trim all continuous-value dependent and predictor variables to the 1st and 99^{th} percentiles. To avoid any bias in assessing the forecast accuracy out-of-sample, there is no such trimming in the data upon which the estimated coefficients are applied to obtain the forecasts. Table 3.1 presents the definitions of all the variables used in our analyses.

TABLE 3.1
Variable definitions

Variable name	Description	Computation / WRDS mnemonic
(USD million)		
OPINC	Operating income after depreciation	OIADP
NI	Income before extraordinary items - available for common equity	IBCOM
TA	Total assets	AT
NOA†	Net operating assets	Common stock (CEQ) + Preferred stock (PSTK) + Long-term debt (DLTT) + Debt in current liabilities (DLC) + Minority interest (MIB) – Cash and short-term investments (CHE)
BV	Common/Ordinary shareholder's equity	CEQ
SALES	Sales/Turnover (net)	SALE
GP	Gross profitability	[Sales (SALE) - Cost of goods sold (COGS)] scaled by Total assets (AT) lagged by one year
OP	Operating profitability	[Gross profit (SALE - COGS) - Selling, general, and administrative expenses reported (XSGA - XRD)] scaled by Total assets (AT) lagged by one year
CbOP_BS‡	Cash-based Operating profitability (balance-sheet approach)	[Operating profit (SALE - COGS - (XSGA -XRD)) - △(Accounts receivable (RECT)) - △(Inventory (INVT)) - △(Prepaid expenses (XPP)) + △(Deferred revenue (DRC+DRLT)) + △(Trade accounts payable (AP)) + △(Accrued expenses (XACC))] scaled by Total assets (AT) lagged by one year
CbOP_CF	Cash-based Operating profitability (cash-flow approach)	[Operating profit (SALE - COGS - (XSGA -XRD)) - Decrease in accounts receivable (RECCH) - Decrease in inventroy (INVCH) - Increase in accounts payable and accrued liabilities (APALCH)] scaled by Total assets (AT) lagged by one year
RNOA	Return on net operating assets	$OPINC_t/(0.5*(NOA_t + NOA_{t-1}))$
ROE	Return on equity	$NI_t/(0.5*(BV_t + BV_{t-1}))$
GSL	Growth in sales	$(SALES_t - SALES_{t-1})/ SALES_{t-1}$

† If the data items for preferred stock, long-term debt, debt in current liabilities, minority interest and cash and short-term investments are not available, they are assumed to equal zero.
‡ If the data items from balance sheet accounts are not available, they are assumed to equal zero.

Panel A of Table 3.2 summarizes the sample selection procedure for the forecasting analysis. The forecasting models are estimated annually on a rolling basis using the previous ten years of data. The actual number of observations used in each industry-specific estimation can vary because only industries with at least 100 firm-year observations are included to avoid unreliable estimation.

TABLE 3.2
Sample selection and descriptive statistics, 1989-2016

Panel A: Sample selection	
Total observations with non-missing SIC code	171,081
Less financial and utility firms, U.S. postal service, and public administration	34,828
Less observations with small denominators	36,420
Less observations with growth exceeding 100%	5,063
Less observations with profitability or growth in sales larger than 1 in absolute value	8,314
Less top- and bottom-percentile observations	5,846
Observations available for in-sample estimation	80,610

This panel summarizes the sample selection procedure and the number of observations available after each filter. The in-sample estimation of the forecasting models is done for each year from 1989 to 2016 on a rolling basis using the previous ten years of data. The actual number of observations used in each industry-specific estimation can vary because only industries with at least 100 firm-year observations are included to avoid unreliable estimation.

Panel B of the table presents the descriptive statistics of the sample with data available for constructing the profitability and sales growth forecasts. On average, the OP and the two versions of CbOP are in the range of 14% to 15.5%, in contrast to the smaller RNOA and ROE (12.6% and 6%, respectively). As GP only has the cost of goods sold deducted, its average value is much higher at 34.2%. The mean growth in sales is 8.2%.

TABLE 3.2 (continued)
Sample selection and descriptive statistics, 1989-2016

Panel B: Descriptive statistics

Variable	Mean	Std. Dev.	Coefficient of Variation	Min.	Median	Max.	Skewness Coefficient	Quartile Skewness
Gross profit	820.2	3,422.3	4.17	-2,865	131.2	130,978	14.86	0.72
Operating profit	487.7	2,074.3	4.25	-4,171	61.4	89,797	14.68	0.77
OPINC	286.9	1,310.8	4.57	-8,715	30.6	71,230	18.69	0.79
NI	140.7	1,010.8	7.18	-98,696	11.3	53,394	3.05	0.82
TA	3,261.4	12,487.9	3.83	10.1	429.4	479,921	13.47	0.78
NOA	2,010.4	7,350.5	3.66	10.0	274.0	281,441	12.64	0.78
BV	1,210.9	4,727.6	3.90	1.1	197.8	174,201	14.64	0.74
SALES	2,673.7	11,526.7	4.31	4.8	431.3	483,521	18.24	0.73
GP	35.1%	21.1%	0.60	-6.8%	31.7%	94.4%	0.66	0.16
OP	15.3%	9.8%	0.64	-19.2%	14.1%	56.9%	0.54	0.26
CbOP_BS	14.5%	10.4%	0.71	-21.8%	13.6%	56.3%	0.32	0.21
CbOP_CF	14.1%	10.1%	0.71	-22.9%	13.3%	54.1%	0.30	0.21
RNOA	12.6%	15.8%	1.26	-65.3%	11.9%	75.7%	-0.10	0.21
ROE	6.0%	17.9%	2.97	-78.4%	9.2%	59.4%	-1.41	-0.36
GSL	7.9%	19.1%	2.42	-55.9%	6.3%	79.3%	0.45	0.20

This panel gives an overview of the data with 80,610 firm-year observations used to compute the forecast improvements for the period from 1989 to 2016. Except for the profitability and growth in sales measures, the descriptive statistics reported are in USD million. *Gross profit* = Sales (SALE) - Cost of goods sold (COGS). *Operating profit* = Gross profit (SALE - COGS) - Selling, general, and administrative expenses reported (XSGA - XRD). See Table 3.1 for other variable definitions. The Coefficient of Variation column reports the ratio of the standard deviation to the mean as a standardized measure of dispersion. The Skewness Coefficient column reports the adjusted Fisher-Pearson standardized moment coefficient of skewness, with negative and positive values representing negative and positive skewness, respectively. The Quartile Skewness column reports the Bowley skewness with a similar interpretation. It is defined in terms of quartiles (Kim and White 2004) and has a value between -1 and +1.

The Coefficient of Variation column of the panel reports the ratio of the standard deviation to the mean. This is a standardized measure of dispersion that can be compared across the profitability measures. The coefficients of variation for the new profitability measures are at very similar levels (from 0.60 to 0.71), whereas those for the traditional measures are substantially higher (1.26 for RNOA and 2.97 for ROE). This suggests that the variations in profitability across firms for a traditional measure are much greater than those for a new profitability measure. Consequently, for a given industry classification, the within-industry heterogeneity in profitability for ROE or RNOA is likely to be greater than that for a new profitability measure. That means, if an industry classification effectively balances the bias-variance tradeoff for forecasting a new profitability measure on an industry-specific basis, the classification is unlikely to do so as effectively for a traditional measure. We will come back to this implication later when discussing the forecasting analysis results.

The Skewness Coefficient column reports the adjusted Fisher-Pearson standardized moment coefficient of skewness. The Quartile Skewness column reports the Bowley skewness defined in terms of quartiles (Kim and White, 2004). All the earnings and size measures are positively skewed (i.e., skewed to the right – with a long, fat right tail). Similarly, most of the profitability and sales growth measures are positively skewed. The only exceptions are RNOA and ROE. Note that they are deflated by the average of the current- and previous-year investment base (i.e. NOA and BV, respectively), unlike the other profitability measures which are deflated by the previous-year total assets (TA). When measured by the moment coefficient of skewness, the negative skewness of RNOA is mild. In contrast, ROE is highly negatively skewed. This is likely due to the relatively light skewness of NI matched to the highly-skewed BV used as the deflator. In terms of the quartile skewness, ROE is still negatively skewed, whereas RNOA is not.

Panel C of the table reports the number of observations and the mean profitability and growth in sales by industry. The Manufacturing sector spreads over two first-digit SIC codes; so does the Services sector.[6] The Agriculture, Forestry, and Fishing sector has the smallest number of observations. There are variations in the industry means of profitability and growth in sales. For the industry GP, it ranges from 17.4% to 49.2%. The variations in the industry OP and CbOP are much smaller (from 11.5% to 17.4%), with the industry RNOA falling into a comparable range (from 9.2% to 14.9%). For the "most polluted" ROE, the industry means vary from 2.5% to 8.5%. The industry GSL has a low at 6.7% and a high at 12.0%.

TABLE 3.2 (continued)
Sample selection and descriptive statistics, 1989-2016

Panel C: Descriptive statistics by industry

1st-digit SIC	Description	Obs.	*GP*	*OP*	*CbOP_BS*	*CbOP_CF*	*RNOA*	*ROE*	*GSL*
0	Agriculture, Forestry, and Fishing	445	26.5%	13.2%	12.1%	12.3%	11.3%	7.0%	7.0%
1	Mining; Construction	5,522	22.0%	14.2%	13.2%	13.0%	9.2%	4.1%	10.9%
2	Manufacturing I	12,517	39.5%	16.3%	15.3%	15.0%	14.9%	8.5%	6.9%
3	Manufacturing II	25,024	39.3%	17.4%	16.2%	15.9%	12.7%	5.2%	7.1%
4	Transportation, Communications, and Sanitary Service	14,380	17.4%	12.8%	12.7%	12.4%	11.4%	8.5%	6.7%
5	Wholesale Trade; Retail Trade	9,946	49.2%	13.2%	11.9%	11.5%	13.5%	6.3%	8.6%
7	Services I	9,365	38.1%	15.8%	16.0%	14.9%	11.5%	2.5%	9.3%
8	Services II	3,411	35.1%	14.3%	13.4%	12.0%	14.0%	4.5%	12.0%
	Overall	80,610	35.1%	15.3%	14.5%	14.1%	12.6%	6.0%	7.9%

This panel reports by industry the number of observations and the mean profitability and sales growth of the sample for the forecasting analysis. Industries are defined using the first-digit Standard Industry Classification (SIC). Financial and utility firms (SIC from 6000 to 6799, or from 4900 to 4949), U.S. postal service (SIC 4311), and public administration (SIC 9000 or above) are excluded from the sample.

[6] In unreported analysis, we consider the industry classification by SIC Division (A to I) with Manufacturing firms and Services firms each under one division and Mining, Construction, Wholesale Trade, and Retail Trade in separate divisions. Although this classification improves the homogeneity of some divisions, it is achieved at the cost of their smaller sizes. Additionally, the classification is likely to worsen the homogeneity of some others by defining them overly broadly. The results based on this industry classification are weaker as expected, consistent with the insight of Schröder and Yim (2017).

To select the sample for the hedge portfolio analysis, we start with all firms traded on the NYSE, Amex, and Nasdaq and retain only observations for ordinary common shares. Following Ball et al. (2015), we use CRSP delisting returns and impute a return of -30% whenever a delisting return is missing and the delisting is performance- related (see also Shumway 1997; Shumway and Warther 1999; Beaver et al. 2007).

We merge the stock data with the accounting data with a lag of six months assuming that accounting disclosure for a fiscal year would be available to the public six months after the fiscal year end. To determine the risk-adjusted returns, we consider both the Carhart (1997) four-factor and the Fama and French (2015) five-factor asset pricing model. We obtain the returns data for the models from Kenneth French's website

(http://mba.tuck.dartmouth.edu/pages/ faculty/ken.french/data_library.html).

3.4. Main Results

3.4.1. Forecasting analysis

Table 3.3 presents the forecasting analysis results comparing the alternative approach by economy-wide quantile regression to the benchmark approach by economy-wide OLS regression. We obtain strong evidence showing significantly positive forecast improvements for all the profitability measures. This holds not only for the mean forecast improvements but also for the median. The levels of significance are consistently high (all at the 1% level). Interestingly, among the new profitability measures or among the traditional ones, the ranking of the magnitudes of the forecast improvements appears to be more or less consistent with the relative degree of skewness of the profitability measures shown in table 2B. Taken together, the findings in table 3 are in line with the expectation that quantile regression can improve the forecast accuracy of profitability measures with skewed distributions.

TABLE 3.3
Profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression

	Value		*p-Value*
GP			
Mean	0.342%	***	0.000
Median	0.379%	***	0.000
OP			
Mean	0.195%	***	0.000
Median	0.196%	***	0.000
CbOP_BS			
Mean	0.112%	***	0.000
Median	0.101%	***	0.000
CbOP_CF			
Mean	0.121%	***	0.000
Median	0.112%	***	0.000
RNOA			
Mean	0.270%	***	0.000
Median	0.235%	***	0.000
ROE			
Mean	0.448%	***	0.000
Median	0.332%	***	0.000

This table reports the profitability forecast improvements of economy-wide quantile regression (the alternative approach) over economy-wide OLS regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same set of predictor variables like those in Fairfield et al (2009). Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see Table 3.1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.4 compares the alternative approach by industry-specific quantile regression to the benchmark approach by economy-wide quantile regression. Again, the mean and median forecast improvements for all the profitability measures are significantly positive. Those for the new profitability measures are highly significant. By contrast, the significance levels of the forecast improvements for ROE are not as high, with those for RNOA following in between.

TABLE 3.4
Profitability forecast improvements of industry-specific quantile regression over economy-wide quantile regression

	Value		*p-Value*
GP			
Mean	0.082%	***	0.000
Median	0.075%	***	0.000
OP			
Mean	0.012%	***	0.002
Median	0.017%	***	0.000
CbOP_BS			
Mean	0.035%	***	0.000
Median	0.044%	***	0.000
CbOP_CF			
Mean	0.019%	***	0.000
Median	0.023%	***	0.000
RNOA			
Mean	0.010%	**	0.048
Median	0.009%	***	0.002
ROE			
Mean	0.011%	**	0.039
Median	0.005%	*	0.064

This table reports the profitability forecast improvements of industry-specific quantile regression (the alternative approach) over economy-wide quantile regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same set of predictor variables like those in Fairfield et al (2009). Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see Table 3.1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

The relative strength of the forecast improvements between RNOA and ROE is similar to prior findings for the industry-specific versus economy-wide comparison under OLS regression (Schröder and Yim 2017). The considerably weaker results for ROE are in line with it being the "most polluted" measures of economic profitability. Interestingly, GP as arguably the cleanest measure of economic profitability also has the largest forecast improvements. The two versions of CbOP, which are found to be better profitability measures than OP in the asset pricing context, also have larger forecast improvements than OP. This suggests that accounting accruals adjustments weaken the linkage of a firm's profitability to its industry membership.

Recall that the variations in profitability across firms are greater for the traditional profitability measures (table 2B). This is likely to result in a higher within-industry heterogeneity in profitability for these measures. The higher heterogeneity would need a not so broad industry classification to effectively balance the bias-variance tradeoff for forecasting the traditional measures on an industry-specific basis. This further explains why in Table 3.4 the forecast improvements for these measures are weaker than those for the new profitability measures.

It is worth noting that the magnitudes of the forecast improvements in Table 3.3 are several to over ten times bigger than those in Table 3.4. This suggests that replacing economy-wide OLS by its quantile regression counterpart is more critical than additionally using the industry-specific version in achieving forecast improvements.

Table 3.5 presents strong evidence showing that forecasts by industry-specific quantile regression are more accurate than their OLS counterparts. Without exception, the mean and median forecast improvements for all the profitability measures are positive and highly significant. We conclude that quantile regression is more accurate than OLS in forecasting profitability, whether on an industry-specific or economy-wide basis.

TABLE 3.5
Profitability forecast improvements of industry-specific quantile regression over industry-specific OLS regression

	Value		*p-Value*
GP			
Mean	0.228%	***	0.000
Median	0.230%	***	0.000
OP			
Mean	0.164%	***	0.000
Median	0.154%	***	0.000
CbOP_BS			
Mean	0.109%	***	0.000
Median	0.101%	***	0.000
CbOP_CF			
Mean	0.106%	***	0.000
Median	0.091%	***	0.000
RNOA			
Mean	0.259%	***	0.000
Median	0.231%	***	0.000
ROE			
Mean	0.428%	***	0.000
Median	0.407%	***	0.000

This table reports the profitability forecast improvements of industry-specific quantile regression (the alternative approach) over industry-specific OLS regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same set of predictor variables like those in Fairfield et al (2009). Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see Table 3.1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

3.4.2. Hedge portfolio analysis

The forecasting analysis results provide clear evidence that quantile regression can improve the accuracy of profitability forecasts over the prevalent approach by OLS regression, regardless of the profitability measures (new or conventional). To assess the economic importance of knowing the quantile regression forecasts in addition to the OLS regression forecasts, we perform a hedge portfolio analysis to see whether that knowledge can lead to abnormal returns.

Table 3.6 presents the average monthly returns of the long and the short portfolio, as well as the average excess returns of the hedge portfolio and the risk-adjusted returns based on the Carhart (1997) four-factor and the Fama and French (2015) five-factor asset pricing model. The hedge portfolio is formed by confining to the firms with positive forecast improvements and sorting them based on the extent the alternative-approach forecasts are higher than the benchmark-approach forecasts. In this table, the alternative forecasting approach is economy-wide quantile regression, whereas the benchmark approach is economy-wide OLS regression. All the average excess returns and the four-factor alphas are significantly positive, except for ROE. The five-factor alphas, however, are all significantly positive, even for ROE. Taken together, the evidence confirms that it is economically important to know the forecasts by economy-wide quantile regression over and above their OLS counterparts.

Table 3.7 shows that forecasts by industry-specific quantile regression also contain economically important information not already found in their OLS counterparts. Like the results in the previous table, the average excess returns and the four-factor alphas are significantly positive for nearly all the profitability measures. The only exception is ROE. Again, the five-factor alphas are all significantly positive, even for ROE. Interestingly, the relative strength of the alphas appears to be by and large in line with (i) the superiority of OP and CbOP over GP in explaining the cross-section of stock returns and (ii) these new profitability measures being more free from financial reporting discretion than the traditional measures. Overall, the results in tables 6 and 7 confirm that the higher accuracy of the profitability forecasts by quantile regression, whether on an industry-specific or economy-wide basis, is useful to investors.

TABLE 3.6
Portfolio returns based on profitability forecasts and forecast improvements: Economy-wide quantile versus economy-wide OLS regression

	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)			Average excess: High − Low			Carhart 4-factor alpha			Fama and French 5-factor alpha		
	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*
GP	0.889%	***	0.000	0.306%		0.233	0.583%	***	0.002	0.513%	***	0.002	0.537%	***	0.001
OP	0.892%	***	0.000	0.258%		0.357	0.634%	***	0.001	0.579%	***	0.000	0.649%	***	0.000
CbOP_BS	0.870%	***	0.001	0.257%		0.361	0.613%	***	0.000	0.660%	***	0.000	0.746%	***	0.000
CbOP_CF	0.898%	***	0.000	0.295%		0.287	0.603%	***	0.000	0.594%	***	0.000	0.596%	***	0.000
RNOA	0.888%	***	0.000	0.518%	*	0.075	0.370%	*	0.052	0.416%	***	0.009	0.331%	**	0.028
ROE	0.881%	***	0.000	0.748%	***	0.004	0.133%		0.375	0.190%		0.145	0.255%	**	0.038

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *DIFF* variable, defined as the alternative-approach forecast in excess of the benchmark-approach forecast, after confining to firms with positive forecast improvements. For this table, the alternative forecasting approach is economy-wide quantile regression, whereas the benchmark approach is economy-wide OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

TABLE 3.7
Portfolio returns based on profitability forecasts and forecast improvements: Industry-specific quantile versus industry-specific OLS regression

	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)			Average excess: High − Low			Carhart 4-factor alpha			Fama and French 5-factor alpha		
	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*
GP	0.894%	***	0.000	0.446%		0.161	0.448%	**	0.025	0.469%	***	0.002	0.571%	***	0.001
OP	0.904%	***	0.000	0.417%		0.153	0.487%	**	0.012	0.468%	***	0.003	0.575%	***	0.001
CbOP_BS	0.886%	***	0.000	0.381%		0.179	0.505%	***	0.005	0.595%	***	0.000	0.641%	***	0.000
CbOP_CF	0.903%	***	0.000	0.235%		0.424	0.667%	***	0.000	0.683%	***	0.000	0.742%	***	0.000
RNOA	0.862%	***	0.000	0.513%	**	0.044	0.348%	**	0.036	0.375%	**	0.016	0.411%	***	0.007
ROE	0.844%	***	0.000	0.651%	***	0.004	0.193%		0.198	0.226%		0.104	0.298%	**	0.028

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *DIFF* variable, defined as the alternative-approach forecast in excess of the benchmark-approach forecast, after confining to firms with positive forecast improvements. For this table, the alternative forecasting approach is industry-specific quantile regression, whereas the benchmark approach is industry-specific OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

3.5. Additional Analyses

We perform additional analyses to ensure that our results are not sensitive to various methodological and sample choices and can extend beyond profitability forecasting.

3.5.1. Forecasting growth in sales

Like the profitability measures, growth in sales also has a skewed distribution (see Table 3.2 Panel B). If skewness is a key reason why forecasts by quantile regression are more accurate than their OLS counterparts, our results for the profitability measures should extend to sales growth forecasting. The findings in Table 3.8 confirm that this is true. Without an exception, the mean and median forecast improvements for growth in sales are positive and highly significant for all the three pairwise comparisons.

3.5.2. Alternative ways to construct PREDGSL

The results in Table 3.8 show that sales growth forecasts by industry-specific quantile-regression are more accurate than their OLS counterparts. In unreported analysis, we re-run the forecasting and hedge portfolio analyses using the *PREDGSL* variable constructed by industry-specific quantile regression. The results are generally consistent.

TABLE 3.8
Sales growth forecast improvements

	Value		*p-Value*
Economy-wide quantile versus economy-wide OLS regression:			
Mean	0.065%	***	0.000
Median	0.095%	***	0.000
Industry-specific versus economy-wide quantile regression:			
Mean	0.025%	***	0.000
Median	0.028%	***	0.000
Industry-specific quantile versus industry-specific OLS regression:			
Mean	0.056%	***	0.000
Median	0.085%	***	0.000

This table reports the sales growth forecast improvements of an alternative approach over a benchmark approach as indicated in the panels of the table. The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same simple first-order autoregressive model specification as in Fairfield et al (2009). Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1979 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for the growth in sales (GSL) measure (see Table 3.1 for the definition of the measure). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

3.5.3. Alternative forecasting model specification

Schröder and Yim (2017) find that the parsimonious first-order autoregressive model (i.e., without the above-median-profitability dummy variable and the PREDGSL variable) can forecast better out-of-sample, even though the Fairfield et al. (2009) specification has a better in-sample estimation fit. In unreported analysis, we re-run our main analyses using the parsimonious specification. The results are generally consistent.

Following Sloan (1996), we use an alternative model by decomposing the profitability (ROA) into accruals and cash flows. This decomposition serves two purposes. First, analysts prefer CFO rather than net income since CFO is less subject to distortion (Bernstein 1993, 461). A higher CFO to net income ratio is believed to have higher earning's quality. Second, a vast literature finds that CFO and accruals have different prediction ability on forecasting profitability (Sloan 1996, Richardson, Sloan, Soliman and Tuna 2005, Joos and Plesko 2005, Konstantinidi and Pope 2016). The less reliable accruals lead to lower earnings persistence (Richardson et al., 2005). We follow the definition of accruals by Sloan (1996):

$$Accruals = (\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD - \Delta TP) - Dep$$

where $\Delta CA = change\ in\ current\ assets\ (Compustat\ item\ 4)$,

$\Delta Cash = change\ in\ cash\ and\ cash\ equivalents\ (Compustat\ item\ 1)$,

$\Delta CL = change\ in\ current\ liabilities\ (Compustat\ item\ 5)$,

$\Delta STD = change\ in\ debt\ included\ in\ current\ liabilities\ (Compustat\ item\ 34)$,

$\Delta TP = change\ in\ income\ taxes\ payable\ (Compustat\ item\ 71)$, and

$Dep = depreciation\ and\ amortization\ expense\ (Compustat\ item\ 14)$.

Earnings used in the forecasting model is ROA which is decomposed by the deflated accrual component and the deflated CFO component.

$$Earnings = \frac{Income\ from\ Continuing\ Operations}{Average\ Total\ Assets}$$

$$Accrual\ Component = \frac{Accruals}{Average\ Total\ Assets}$$

$$Cash\ Flow\ Component = \frac{Income\ from\ Continuing\ Operations - Accruals}{Average\ total\ Assets}$$

Other literature documents asymmetric earnings persistence due to the binary classification of firms into profits and losses. Hayn (1995) posits that losses are less persistent than profit due to the exercising of abandonment options by firms. Basu (1997) argues that the timely recognition difference between losses and profits causes the persistence variation where earnings reflects bad news more quickly than good news. To capture any difference in profitable and loss-making firms, we include a dummy variable LOSS_D which equals one if the firm's ROA is negative and zero otherwise.

Our forecasting approach uses the following earnings decomposition model:

$$ROA_{i,t} = \alpha_{j,T} + \beta_{j,T}ACCRUALS_{i,t-1} + \gamma_{j,T}CFO_{i,t-1} + \lambda_{j,T}LOSS\_D_{i,t-1} + \varepsilon_{i,t}, \qquad (11)$$

The results of the forecasting improvement using this decomposition method are present in Appendix 3.1. We find significant and positive forecasting improvement of using quantile regression over OLS regression for all the model specifications.

3.5.4. Alternative industry classifications

In unreported work, we re-run our main analyses using alternative industry classifications, such as Fama-French 12-industry and two-digit SIC. Fama-French 12-industry is a broad industry classification similar to the first-digit SIC. The results for this industry classification are generally consistent and sometimes even stronger. Because the two-digit SIC is a narrower classification, the results are weaker though qualitatively similar. This is in line with the insight of prior research (Schröder and Yim 2017).

3.5.5. Sample period before financial crisis

In unreported analysis, we re-run the forecasting analysis for the reduced sample up to 2006 (i.e., before the financial crisis). The results are very similar, though occasionally weaker. This is likely to be have been driven by the sharp reduction in the sample size compared to the full sample from 1989 to 2016.

3.5.6. Long-term forecasting analysis

We also perform a series of forecasting accuracy comparison between quantile regression and OLS regression in predicting long-term forecasts. In Appendix 3.2, we present the forecasting improvement of predicted profitability by using quantile regression over OLS regression under the economy-wide model. From 2-year ahead forecasts to 5-year ahead forecasts, the results show quantile regression forecasts as more accurate than OLS regression forecasts. This finding is robust for all our profitability measures. In further unreported results, we confirm this finding using the industry-specific forecasting model.

3.5.7. Robustness tests of hedged portfolio analyses

In Section ***3.4.2*** we show the economic importance of quantile regression forecasts over OLS regression forecasts in terms of obtaining abnormal return through hedged portfolio analyses. A hedged portfolio is formed by confining to firms with positive forecast improvements and sorting them based on the extent the alternative-approach forecasts are higher than the benchmark-approach forecasts. If market participants trade or price stocks based on forecasts of firm profitability obtained with OLS, stock price would not fully impound the information contained in quantile regression forecasts. We value this informative of quantile regression as the difference of predicted profitability based on quantile regression and OLS regression (DIFF). As we did in Table 3.6 and Table 3.7, we sort stocks based on DIFF with stocks having positive forecasting improvement of using quantile regression forecasts over OLS regression forecasts. This method is naturally biased to the quantile regression method as we only include the firms which have more accurate forecasts by quantile regression. In this section, we provide two additional tests based on two ways to confine the stock sample into portfolio to further confirm the usefulness of quantile regression forecasting in hedged portfolio analyses.

In the first set of analyses, we construct a trading strategy based on DIFF with a sample of firms whose forecasting improvement of quantile regression forecasts over OLS regression forecasts is negative. This method of strategy is in favour of OLS regression forecasts since we only include the firms which have more accurate predicted profitability by using OLS regression over quantile regression. In Appendix 3.3 Panel A and B, for both economy-wide and industry-specific forecasting models, we cannot obtain any positive and significant abnormal returns by sorting stocks based on the negative forecasting improvement. This is consistent with fact that OLS regression is widely used by the market participants in trading their stocks. Therefore, no further positive abnormal returns can be generated based on the advantages of OLS regression forecasts.

In the second set of analyses, we construct the trading strategy based on DIFF with the full sample regardless of whether the forecasting improvement is positive or negative by using quantile regression over OLS regression in forecasting. In Appendix 3.4 Panel A and Panel B, we find that for either economy-wide or industry-specific model, our trading strategy can obtain positive and significant abnormal returns for all our profitability measures.

These two additional trading strategies further confirm the usefulness of quantile regression forecasts over OSL forecasts. Trading strategy guided by quantile regression forecasts when quantile regression is more accurate brings a higher abnormal return than trading guided by OLS forecasts when OLS forecasting is more accurate.

3.6. The implication of quantile regression in analysts' forecasts

In this section, we conduct several tests by comparing our model-based forecasting model and analysts' consensus forecasts. Though analysts' forecasts are widely explored by prior research, few studies compare the forecasts between analysts and forecasting model directly.

First, analysts' forecasts have low coverage of data compared to model-based forecasts. We collect analysts' forecasts from I/B/E/S file. The figure 3.2 shows a coverage ratio by different earing's measures per year from I/B/E/S and Compustat. The earning's measures we compare are non-deflated earnings which are gross profit, operating profit, and EPS, and deflated earnings which are ROE, and ROA. The figure suggests that all the profitability measures from Compustat maintain a consistent level of coverage from 1960s to 2017. By contrast, all profitability measures except EPS from I/B/E/S have very low coverage and only available after 2000. The un-deflated EPS from I/B/E/S have a consistent trend of number of observations with the profitability measures in Compustat since early 1980s though the coverage is lower.

Second, there is an inconsistency of data between these two databases. The actual and analysts' forecasts from I/B/E/S are adjusted with non-recurring items while the earnings from Compsutat are not which follows GAAP.

Therefore, we reconcile these differences of the earnings between these two databases by making the following adjustments. We use the IBES actual EPS as our main earnings in our model-based forecast model. To convert EPS into deflated profitability ROE, we use two alternative deflators.

Method1: $IBES\_EPS\_P_t = IBES\_EPS_t / PRICE_{t-2}$

Method2: $IBES\_EPS\_CH\_SEQ_t = (IBES_{EPS_t} * Csho_{t-1}) / Avg.SEQ_{t-1}$

where $PRICE_{t-2}$ is the fiscal year end price obtained from I/B/E/S. We follow Basu and Markov (2004) by using the lagged two year's price as the deflator. CSHO are common shares outstanding and avg. SEQ are the average book value of equity from Compustat. For the following sections, these two ways of conversion of ROE are applied to the comparison between analysts' forecasts and model-based forecasts.

Figure 3.2

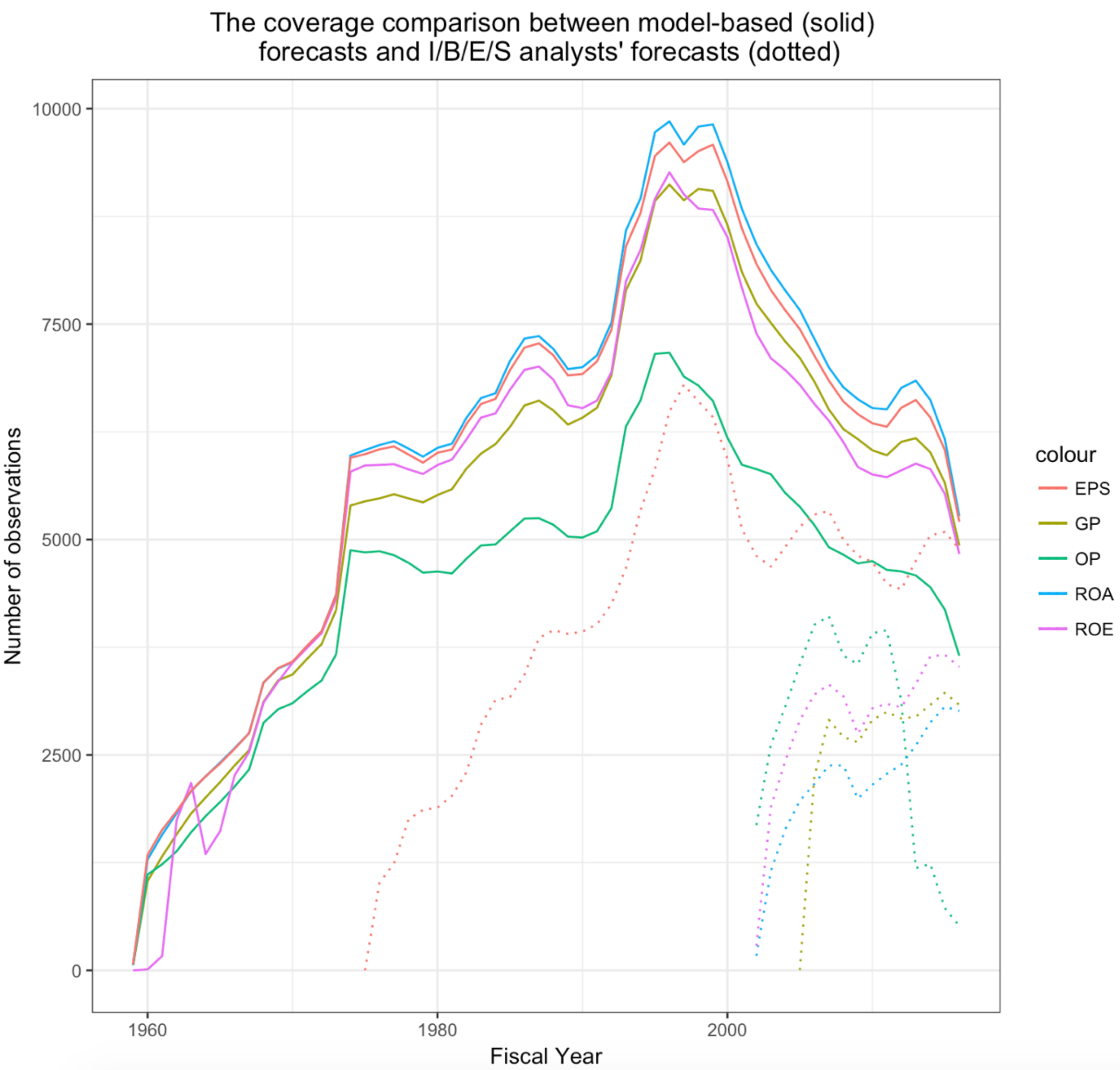

Figure 3.2 summarises the coverage of earnings from Compustat database and I/B/E/S database from 1960s to 2016. Earnings per share (EPS), gross profit (GP) and operating profit (OP) are the non-deflated earnings. ROA and ROE are profitability ratios collected from I/B/E/S directly, while for Compustat database we calculate ROA and ROE following their definitions in Table 3.1.

3.6.1. Model-based forecasts and analysts' forecasts accuracy comparison

In this section, we compared the forecasting accuracy between analysts' forecasts and the model-based forecasts. We use the economy-wide parsimonious forecasting model to construct our model-based forecasts.

$$x_{i,t+n} = \alpha_T + \beta_T x_{i,t} + u_{i,t}, \qquad (12)$$

where x is the ROE defined as IBES_EPS_P or IBES_EPS_CH_SEQ. *i* is the firm and *t* is the fiscal year. We conduct 1 year ahead to 5 years ahead forecasts as well as quarter 1 to quarter 4 ahead forecasts in order to give a comprehensive comparison. Therefore, *n* is 1 year to 5 year or quarter 1 to quarter 4.

Like our forecasting procedure in the main tests of this paper, we estimate the in-sample coefficients using either OLS regression or quantile regression in the rolling basis with the past 10 years of data. The profitability forecasts using the model is predicted as follow:

$$E_{\text{ew\_OLS or QR}}[x_{i,T+n}] = a_T + b_T x_{i,T},$$

where (a_T, b_T) are the OLS or quantile regression estimates of the economy-wide model parameters (α_T, β_T).

For a fair comparison, we construct analysts' forecasts from I/B/E/S detail file and restricts the analysts' forecasts after the most recent annual earnings announcement date but before the first quarter's earnings announcement date for the following year. During this time interval, both analysts' forecasts and model-based forecasts have the same information of the firm's prior performance. For example, if the forecasted variable is $x_{i,t+1}$, we aggregate all the analysts' forecasts for *t+1* during the period between *t* and *t+ Q1(the quarter 1's earnings announcement date)*. We divide analysts EPS forecasts with our two methods of deflators.

Method1: $IBES\_Analysts\_EPS\_P_{t+n} = IBES\_Analaysts\_EPS_{t+n}/PRICE_{t-2}$

Method 2: $IBES\_Analysts\_EPS\_CH\_SEQ_{t+n} = (IBES\_Analysts\_EPS_{t+n} * Csho_t)/avg.SEQ_{t-1}$

We use the mean and median analysts' consensus forecasts in the comparison with the model-based forecasts. Consistent to our earlier analyses, we use the absolute forecast error (AFE) to measure the accuracy of a forecasting approach.

We summarise the data descriptive analyses of IBES profitability (ROE) based on our deflators (Method 1 and Method 2) and the corresponding deflated analysts' consensus forecasts in the Appendix 3.5 Panel A and Panel B. Since EPS in I/B/E/S only has a good coverage after 1980, we select our sample period from 1985 to 2017 to ensure enough observations for in-sample estimation. Similar to the Compustat-based ROE reported in Table 3.2 Panel B, both annual and quarterly IBES-based ROE constructed by both deflators are negatively skewed. IBES_EPS_P is smaller than IBES_EPS_CH_SEQ due to the different deflating measures we used.

We present the profitability forecasting accuracy of the model-based and analysts' forecasts by using the deflator proposed in Method 1 for both actual profitability and analysts' forecasts' consensus. In Table 3.9 Panel A, we present the forecasting accuracy comparison between analysts' consensus forecasts and our model-based forecasts in forecasting quarter 1 to quarter 4 ahead quarterly earnings. We find that for quarter 1 to quarter 4 earnings' forecasts, analysts give more accurate forecasts (both mean and median consensus) than our model-based forecasts either by using OLS regression or quantile regression.

We repeat the same forecasting process by using the long-term annual forecasts. In Table 3.9 Panel B, we report the forecasting accuracy comparison from 2 years ahead to 5 years ahead forecasts by using our two deflated earnings. We find that analysts' forecasts are only more accurate than model-based forecasts in the 1 year ahead forecasts but less accurate from 2 years ahead onwards.

For the 'horseracing' comparison between the two model-based forecasts, quantile regression forecasts are more accurate than OLS forecasts from quarter1 to quarter 4 and 1 year ahead to 5 years ahead. The same findings are shown by using the deflators proposed in Method 2 (results are documented in Appendix 3.6 Panel A and B). Overall, our results suggest that analysts' forecasts produce more accurate earnings forecasts in the short term but model-based forecasts are more accurate in predicting long-term earnings.

TABLE 3.9
Profitability forecast improvements of quantile regression over OLS regression by economy-wide model (Column approach forecasts versus row approach forecasts): IBES-based ROE by Method 2

Panel A: Quarterly forecasting

Q1	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.620%	***	0.000	-0.624%	***	0.000	0.015%	***	0.000
Median	-0.362%	***	0.000	-0.366%	***	0.000	0.018%	***	0.000
PRED_OLS									
Mean	-0.634%	***	0.000	-0.639%	***	0.000			
Median	-0.392%	***	0.000	-0.394%	***	0.000			

Q2	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.491%	***	0.000	-0.489%	***	0.000	0.023%	***	0.000
Median	-0.277%	***	0.000	-0.277%	***	0.000	0.029%	***	0.000
PRED_OLS									
Mean	-0.514%	***	0.000	-0.512%	***	0.000			
Median	-0.317%	***	0.000	-0.318%	***	0.000			

Q3	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.399%	***	0.000	-0.399%	***	0.000	0.019%	***	0.000
Median	-0.224%	***	0.000	-0.223%	***	0.000	0.027%	***	0.000
PRED_OLS									
Mean	-0.418%	***	0.000	-0.418%	***	0.000			
Median	-0.261%	***	0.000	-0.261%	***	0.000			

Q4	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.327%	***	0.000	-0.330%	***	0.000	0.013%	***	0.000
Median	-0.193%	***	0.000	-0.194%	***	0.000	0.016%	***	0.000
PRED_OLS									
Mean	-0.341%	***	0.000	-0.343%	***	0.000			
Median	-0.214%	***	0.000	-0.216%	***	0.000			

This table reports the short-term quarterly profitability forecast improvements of economy-wide model-based forecast model (the alternative approach by row) over analysts' consensus forecasts (the benchmark approach by column). The actual earnings and analysts' forecasts are deflated by lagged two years' fiscal closing price (Method1). See the detailed definitions of the variables in the main texts. The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches which equals the AFE by using the column approach minus the AFE by using the row approach. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both model-based forecasting approaches use the same forecasting steps as those in Fairfield et al (2009) but simply based on a parsimonious AR (1) model. Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see main texts for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 3.9
Profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression (Column approach forecasts versus row approach forecasts): IBES-based ROE by Method 2

Panel B: Long-term forecasting

Y1	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-1.211%	***	0.000	-1.210%	***	0.000	0.184%	***	0.000
Median	-0.666%	***	0.000	-0.661%	***	0.000	0.226%	***	0.000
PRED_OLS									
Mean	-1.395%	***	0.000	-1.394%	***	0.000			
Median	-0.933%	***	0.000	-0.934%	***	0.000			

Y2	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	0.329%	***	0.000	0.334%	***	0.000	0.207%	***	0.000
Median	0.220%	***	0.000	0.222%	***	0.000	0.154%	***	0.000
PRED_OLS									
Mean	0.212%	***	0.000	0.123%	***	0.000			
Median	0.093%	***	0.000	0.096%	***	0.000			

Y3	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	1.107%	***	0.000	1.092%	***	0.000	0.133%	***	0.000
Median	0.480%	***	0.000	0.477%	***	0.000	0.136%	***	0.000
PRED_OLS									
Mean	0.974%	***	0.000	0.959%	***	0.000			
Median	0.370%	***	0.000	0.366%	***	0.000			

Y4	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	1.722%	***	0.000	1.677%	***	0.000	0.175%	***	0.000
Median	0.710%	***	0.000	0.699%	***	0.000	0.171%	***	0.000
PRED_OLS									
Mean	1.547%	***	0.000	1.503%	***	0.000			
Median	0.601%	***	0.000	0.587%	***	0.000			

Y5	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	2.333%	***	0.000	2.368%	***	0.000	0.305%	***	0.000
Median	0.824%	***	0.000	0.834%	***	0.000		***	0.000
PRED_OLS									
Mean	2.028%	***	0.000	2.063%	***	0.000	0.311%		
Median	0.615%	***	0.000	0.629%	***	0.000			

This table reports the long-term annual profitability forecast improvements of economy-wide model-based forecast model (the alternative approach by row) over analysts' consensus forecasts (the benchmark approach by column). The actual earnings and analysts' forecasts are deflated by lagged two years' fiscal closing price (Method1). See the detailed definitions of the variables in the main texts. The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches which equals the AFE by using the column approach minus the AFE by using the row approach. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both model-based forecasting approaches use the same forecasting steps as those in Fairfield et al (2009) but simply based on a parsimonious AR (1) model. Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see main texts for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

3.7. Conclusion

Prior research has examined the properties of earnings and profitability and alternative ways to forecast them (Brown and Ball 1967; Brooks and Buckmaster 1976; Freeman et al. 1982; Penman 1991; Rumelt 1991; Lipe and Kormendi 1994; Fama and French 2000; Cheng 2005; Fairfield et al. 2009; Li 2011; Li et al. 2014; Schröder and Yim 2017; Chang et al. 2016). Despite the variations considered, all use OLS regression to construct forecasts.

In this study, we explore two ways to obtain more accurate point forecasts of profitability and assess them against the traditional approach. First, we use quantile regression to construct forecasts, as opposed to the prevalent method by ordinary least squares (OLS) regression. Second, forecasts are constructed on an industry-specific basis, as opposed to the common practice of constructing forecasts on an economy-wide basis. We obtain strong evidence that quantile regression produces more accurate points forecasts of profitability than OLS regression for a number of new and traditional profitability measures, whether on an economy-wide or industry-specific basis.

To assess the economic importance of using the more accurate profitability forecasts by quantile regression, we perform a hedge portfolio analysis similar to that of Li et al. (2014) to link forecast accuracy with stock return predictability. A hedge portfolio is formed by sorting stocks according to the excess of the quantile regression forecast over its OLS counterpart, confining to those stocks with the quantile regression forecasts proven to be more accurate. The higher accuracy suggests that guidance by the quantile regression forecasts rather than by their OLS counterparts should be followed. Therefore, we go long the stocks with the highest excess and short those with the lowest. We find that, regardless of the profitability measure, the hedge portfolio has a significantly positive risk-adjusted return. This suggests that the quantile regression forecasts contain economically important information not already contained in their OLS counterparts.

We contribute to the literature by examining the accuracy and usefulness of forecasting profitability by industry-specific quantile regression. This is found to be the most accurate approach, compared to its economy-wide counterpart and the industry-specific and economy-wide OLS regression approaches. Investors can economically benefit from the information contained in quantile regression forecasts of profitability for a number of new and traditional profitability measures. Our forecasting and hedge portfolio analysis results are not sensitive to various methodological and sample choices.

Figure 3.1

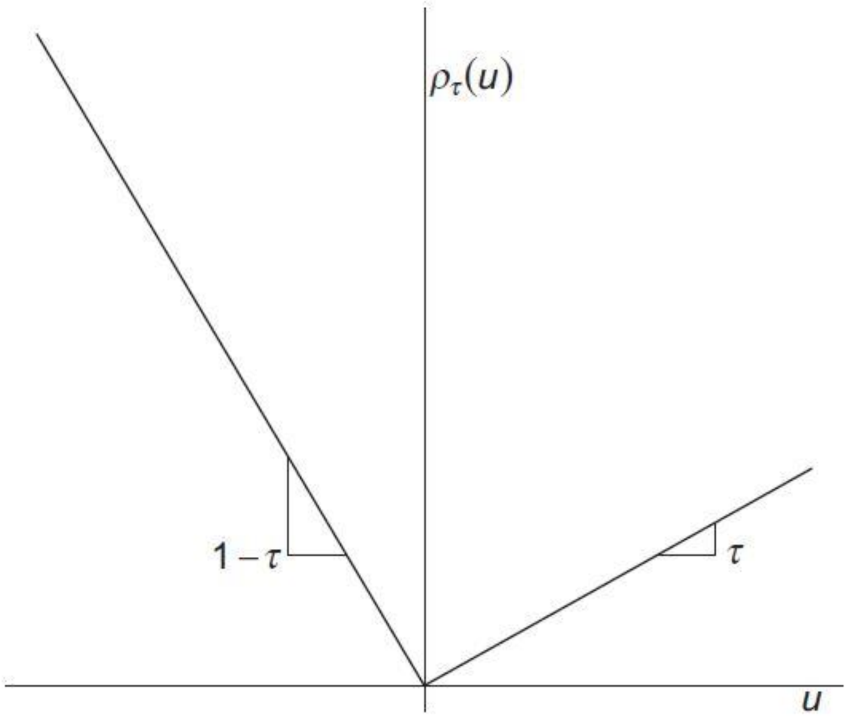

Figure 1 Quantile regression ρ function.

Source: Koenker (2015)

Chapter 4

Firm Diversification and Loss Reversal Probabilities: Evidence from Abandonment Options

4.1. Introduction

The percentage of U.S. firms reporting losses, as recorded on Compustat, increased from 14% in 1979 to 47% in 2016. It is a long-established observation that losses are less persistent than profits because they can be avoided through abandoning loss-making assets (Hayn 1995). Interestingly, the proportion of firms reporting losses is much smaller for diversified firms than for focused firms (see Figure 4.1). Empirical studies suggest that diversified firms differ from focused firms: diversified firms can more easily sell their assets to raise funds (Shleifer and Vishny 1992, Subramaniam et al 2011), they have better credit ratings giving access to cheaper external funds (Lewellen 1971, Dimitrov and Tice 2006), and they have more efficient internal fund transfer channels (Stein 1997, Lamont 1997, Shin and Stulz 1998, Khanna and Tice 2001, Subramaniam et al 2011). In contrast, focused firms tend to have better investment opportunities and are more likely to engage in positive net present value projects (Wernerfelt and Montgomery 1988, Lang and Stulz 1994).

In this chapter, we investigate the role of firm structure in explaining differences in loss reversal probabilities across firms, analysing the liquidation of underperforming assets, segments or entire firms as real options. Basic economic theory suggests that shutting down the business and exit from the market to stem losses is always a natural solution (Samuelson 1948). Further explanation via the abandonment option also has a long history in the literature (Berger, Ofek and Swary 1994, Hayn 1995) but there is less certainty on how to define and quantify any such abandonment option. Inspired by Lawrence, Sloan, and Sun (2017) and their proxy for curtailment, we use this as a key factor to capture different managerial behaviour between diversified firms and focused firms when firm-level losses occur. By adding an abandonment option into the loss reversal model of Joos and Plesko (2005), we treat the option not only as a tool to limit loss but also as part of a strategic view of firms' loss avoidance behaviour.

Our findings are threefold. First, we find that there is a difference in loss reversal probability between diversified firms and focused firms. The higher the diversity of the firm, the higher the probability of loss reversal in the following year. Second, diversified firms can exercise their options to liquidate unprofitable segments more efficiently to stem future losses than focused firms. Third, over-investment as a proxy of the agency problem dampens the efficiency of diversified firms in exercising their real options.

This chapter contributes to the literature in three strands. First, our approach is based on Hayn's (1995) abandonment option hypothesis that shareholders of loss firms will liquidate or curtail the assets of the firm if they expect the losses to continue. Several studies follow and test her theory by quantifying real actions in exercising the abandonment options (Pinnuck and Lillis 2007, Lawrence et al., 2017). We close the unsolved question posed by Pinnuck and Lillis on whether abandonment options result in improved performance, between diversified firms and focused firms in our case, by using a loss reversal model.

Second, we extend the loss reversal probability forecasting model of Joos and Plesko (2005) by involving a firm's diversity level as a key variable. Prior literature in earnings persistence simply pools all the firms without distinguishing between firm structure differences. Our analysis confirms the importance of firm structure in explaining loss reversal capability due to the higher efficiency of exercising the abandonment options by diversified firms.

Our final contribution is in giving support to the firm diversification discount literature. Diversification has two sides to its effects on firms. On the one hand, diversified firms can enjoy the benefit of coinsurance from their multiple segments. Firm-level risk can be diversified through imperfectly-correlated segments' business operations. Therefore, diversified firms can benefit from higher debt capacity and an interest tax shield (Lewellen

1971, Majd and Myers 1987, Hann et al., 2013). In addition, the larger internal capital market of the diversified firms may reduce the problem of underinvestment (Stultz 1990). Our findings are relevant to the diversification premium literature since diversified firms can liquidate their under-performing assets more efficiently than focused firms in order to achieve loss reversal. In addition, diversified firms can more likely sell their assets than focused firms (Subramaniam et al. 2011). Such loss-curtailing actions is value-enhancing as firms can use this low cost internal financing to invest in better profit-making projects. On the other hand, diversification can be costly, mainly due to the agency problem. Managers of diversified firms are more likely to pursue resources for personal empire-building (Jenson 1986), power grabbing (Rajan, Servaes, and Zingales 2000), or through weaker managerial incentives to maximize shareholder value (Denis, Denis, and Sarin 1997). Consequently, over-investment is more severe in diversified firms; discretionary investments in unprofitable projects dampen the firms' performance. Our results further support the agency literature by showing that the superior efficiency in exercising the abandonment option by diversified firms is destroyed by agency problems. This is consistent with the finding that agency problems determine managers' reluctance towards divestment of assets (Boot 1992, Pinnuck and Lillis 2007).

The remainder of the chapter proceeds as follows. Section 2 describes our main hypotheses concerning loss reversal and firm diversification. Section 3 introduces the data and in Section 4 we present our models. Section 5 presents our main results. We provide three econometric methodologies to solve the endogeneity problems in Section 6 and additional robustness tests in Section 7. Section 8 concludes.

4.2. Hypothesis Development

The motivation of the chapter proceeds in two stages. First, we seek to capture the importance of a firm's structure in driving loss reversal probability. Second, we apply abandonment option theory to explain the loss reversal capability difference between diversified firms and focused firms. We conduct four sets of tests. The first set seeks to investigate whether there is any difference of loss reversal ability between diversified firms and focused firms. In the second set, we use abandonment option theory to explain the loss reversal phenomenon. Because diversified firms hold more different business assets than focused firms, we expect that a firm's diversity will lead to the exercise the abandonment option and probable loss reversal. Our third set of tests focus on the cost effects of a firms' diversity. Since diversified firms more easily suffer agency problems, this could reduce the efficiency of abandonment option exercise. In the last set of tests, to control for endogeneity, we present several robustness tests.

Firms are loss-avoiding by nature (Burgstahler and Dichev 1997, Degeorge et al. 1999; Graham et al. 2005). They tend to report higher earnings for an implicit warranty to charge higher prices, for better deals from suppliers and lenders and for keeping valuable employees (Bowen et al., 1995). The classic theory from Hayn (1995) is related to firms' loss avoidance by supposing that shareholders of loss-making firms can always liquidate the firm rather than suffer continuing losses. She predicts that, because of the abandonment option, loss-making firms have lower or insignificant earnings response coefficient (ERC) and R^2 comparing to positively significantly profitable firms. Subsequent studies support Hayn's conclusions that the abandonment option explains the weak ERC of loss firms (Berger et al. 1996, Subramanyam and Wild 1996). Joos and Plesko (2005) extend Hayn's work by proposing a logistic model to estimate loss reversal probability using concurrent and past financial information. Based on the estimated probability of reversal, a firm's loss can be grouped as transitory with the highest estimated probabilities of loss reversal and persistent loss with the lowest likelihood of loss reversal. They predict that if persistent loss indicates a high likelihood of abandonment, the ERC will be insignificant or lower than in the transitory loss groups. In other words, investors predict that persistent loss is not informative about the future performance of firms since exercising the abandonment option can avoid further losses. Conversely, if transitory losses indicate a low likelihood of exercising the abandonment option, this can have a positive effect on stock returns. Li (2011) uses a similar profit forecasting model. Using the framework of Mishkin (1983), Li finds that investors do not fully distinguish the difference in persistence of loss and they tend to treat all losses as transitory.

Building on the loss reversal model of Joos and Plesko (2005), we conduct our primary analysis by testing whether there is any difference in loss reversal ability between these two types of firms when suffering loss, giving our first hypothesis:

Hypothesis (H1). Diversified firms and focused firms have different loss reversal probabilities and that difference is reflected through each firm's diversity level.

Our second hypothesis concerns the relation between the abandonment option and asymmetric loss reversal by interacting the abandonment option and a firm's diversity. The existence of curtailment is difficult to capture because of the unavailability of data and relevant information disclosure. Prior research uses abandonment option theory to explain why losses are less persistent than profit (Hayn 1995, Basu 1997); however, there is no direct measure to identify the presence of curtailments. Pinnuck and Lillis (2007) show that divestment of a division can be reflected in a cut of employees. They find that a firm's reporting of an accounting loss acts as a heuristic trigger for the exercise of an abandonment option. Lawrence, Sloan and Sun (2017) identify a curtailment as when the sales and employee numbers are both less than in the previous period. We borrow their method in our main analysis and in the robustness analysis we use two additional methods to support our findings.

Diversified firms have different business segments while focused firms are only in one line of business activity. Therefore, diversified firms hold a selection of abandonment options through different segments with imperfectly correlated cash flows. When a loss occurs, they can liquidate unprofitable business segments and keep the other businesses operating normally.

Conversely, focused firms hold only one abandonment option for their single type of business. Thus, focused firms are unlikely to abandon their business easily, since exercising that option means shutting down the firm forever. If focused firms determine upon entering a new business, it is difficult to make profits in the short run. Even though part of the loss-making assets can be liquidated, the normal operating business activities can still be damaged severely. Therefore, we hypothesise the following:

Hypothesis 2.1 (H2.1). Among loss firms, diversified firms have a higher frequency of exercising the abandonment option than focused firms.

Hypothesis 2.2 (H2.2). Diversified firms have a higher probability of having losses reversed by exercising their abandonment options.

There is a trade-off in exercising abandonment options for loss-making firms. The benefit of liquidating unprofitable assets or segments is the curtailment of loss and pursuit of profit. Firms are profit-seeking not only because those with consistent profits can build their reputations but also because profitable firms can access external funding at lower cost. However, exercising the abandonment option is irreversible and, importantly, it is part of the firm's strategic plan as determined by its managerial decisions. The agency problem is value-destroying for the firms (Jensen 1986, Stulz 1990) because managers derive private benefits from diversification which is more than their private costs (Denis, Denis and Sarin 1997). In other words, managers may maintain an under-performing diversification strategy even though doing so reduces the benefit to shareholders. Boot (1992) shows that managers may choose to avoid a value-maximizing divestiture because doing so may adversely affect perceptions of their abilities and reputations. In addition, Chen and Rey (2012) suggest that cross-subsidization arises when a firm is engaged with high market power in some markets and can use their more protected revenue to finance losses in a competitive market, known as its "deep pocket". Thus, diversified firms are more likely to finance loss-making segments from their cash rich segments instead of exercising the abandonment options.

We are interested in how such agency problems influence the efficiency of exercising the abandonment options. We use over-investment as a proxy for the agency problem since it is a potential source of value loss from diversification. We measure a firm's overinvestment as whether Tobin's Q is below the industry's median (Berger and Ofek 1995, Subramaniam, Tang, Yue, and Zhou 2011). Jensen (1986, 1993) points out that managers tend to engage in overinvestment out of free cash flow. Scharfstein and Stein (2000) find that less profitable divisions tend to be subsidized by segments with high profitability through the diversified firms' internal capital market. Firms may be reluctant to abandon under-performing business segments because doing so can be harmful to a manager's reputation and business competence (Jensen 1986, Boot 1992). We expect that diversified firms suffers severer agency problems than focused firms. If such agency problems exist, the headquarters (CEO) of diversified firms can play with more resources from the conglomerates or misallocate more resources to the underperforming segments rather than curtailing these segments or delay abandoning them until after the optimal time. Therefore, we propose the following:

Hypothesis 3 (H3). The efficiency of diversified firms, when they are engaged in over-investment, is dampened severely which makes it harder for focused firms to get loss reversal through the abandonment option.

4.3. Data and Sample Selection

We collect data from three sources via Wharton Research Data Services (WRDS). The firm level and business segment data are obtained from Compustat North America annual fundamentals file and segment files. Stock returns data used in portfolio constructions come from the Centre for Research in Security Prices (CRSP) monthly stock file. We sample for the loss reversal model from 1979 to 2016. Firms were required to report segment information after December 15, 1977 by Statement of Financial Accounting Standard (SFAS) No.14. Since our empirical test involves sales growth forecasting requiring past five-years sales data to construct curtailment option, we use firm level data for forecasting as early as 1974.

We exclude financial firms (SIC from 600 to 6799) from our sample. Segment data are merged with firm level data to construct our matched samples. We remove any mismatched observations which have summation of segment sales (segment asset) more than 101% or less than 99% (more than 125% or less than 75%) of the firm's total sales (total assets). A detailed description of our variables is provided in Table 4.1. For all the balance sheet data constructing accruals, we replace the missing values with zero following Ball et al. (2016). For all our empirical tests, we trim all continuous-value variables to the 1st and 99th percentiles.

Table 4.1
Variable definitions

Variable name	Description	Computation / WRDS mnemonic
(USD million)		
IB	Income before extraordinary items	IB
TA	Total assets	AT
NI	Net income	NI
DIVD	Total dividend	DIV
SALES	Sales/Turnover (net)	SALE
ROA	Return on asset	Income before extraordinary item (IB) scaled by Total assets (AT) lagged by one year
MV	Market value of firm	Price_Fiscal Year_Close (PRCC_F) * Common shares outstanding (CSHO)
SIZE	Firm size	Log of MV
SPI	Special items	Special items (SPI) scaled by Total asset (AT) lagged by one year
PAST_ROA	Average return on asset in the past five years	[ROA(t-1) + ROA(t-2) +……+ ROA(t-5)]/5
LEVERAGE	Leverage level	[Long- term debt (DLTT) + Current liabilities (LCT)] / [Long-term debt (DLTT) + Current liabilities (LCT) + Market value (MV)]
T_Q	Tobin's Q	Market value (MV) / Total assets (TA) lagged by one year
GSL	Growth in sales	$(SALES_t - SALES_{t-1}) / SALES_{t-1}$
HERF(TA)	Herfindahl index by total asset	$Herfindahl\ Index, H = \Sigma_{i=1}^{Njt}(\frac{x_{ijt}}{\Sigma_{i=1}^{Njt} X_{ijt}}^{2})$
HERF(SALE)	Herfindahl index by sales	
DIVERSITY_D	Diversity dummy variable	*Equal to one if firm has more than one business segments, zero otherwise*
FIRST_LOSS	First loss dummy variable	*Equal to one if current year is not the first year of loss (ie., last year was profitable), zero otherwise*
LOSS_SEQ	Loss sequence	*The number of loss sequence in the past five years*
DIVDUM	Dividend paying dummy variable	*Equal to one if paying dividend in the current loss year, zero otherwise*
SIC2_MED_T_Q	Median Tobin's Q in yearly basis by 2-digit Standard Industry Classification	*Median Tobin's Q by year and 2-digit Standard Industry Classification*
AB_D	Abandonment option activities dummy variable	*Equal to one if both sales and number of employees decrease compared to previous year, zero otherwise*

† If the data items for preferred stock, long-term debt, debt in current liabilities, minority interest and cash and short-term investments are not available, they are assumed to equal zero.
‡ If the data items from balance sheet accounts are not available, they are assumed to equal zero.

4.4. The earnings forecast model for loss firms

Building on the models of Joos and Plesko (2005) and Li (2011), we develop our annual earnings forecast models for loss firms in the following two specifications:

$$y_{t+1} = \alpha + \beta_1 EARN_t + \beta_2 PAST\_EARN_t + \beta_3 SIZE_t + \beta_4 SALESG_t + \beta_5 FIRSTLOSS_t + \beta_6 LOSS\_SEQ_t + \beta_7 DIVDUM_t + \beta_8 SPI_t + \beta_9 DIVER_t + \beta_{10} LEVERGE_t + \beta_{11} CAPX_t + \beta_{12} T\_Q_t + \beta_{13} CR_t + \varepsilon_{t+1} \quad (1)$$

where our dependent variable y_{t+1} is $REVERSAL_{t+1}$. $REVERSAL_{t+1}$ is an indicator that is one if the loss firm becomes profitable in the next year and zero otherwise. $PAST\_EARN_t$ is defined as the average of past five-year ROA. $EARN_t$ is the firm's profitability measurement using return on asset (ROA). It is defined as annual income before extraordinary items and discontinued operations (annual Compustat item #18) deflated by lagged total asset (annual Compustat item #6). $SIZE_t$ is the logarithm of market value of equity (annual Compustat data item#199*annual Compustat data item#25). $SALESG_t$ is the percentage growth in sales over year *t*; $FIRSTLOSS_t$ is an indicator variable that is equal to one if the current loss is the first in a sequence and zero otherwise; $LOSS\_SEQ_t$ is the number of sequential annual loss over the past five years; $DIVDUM_t$ [7]is an indicator variable that is equal to one if the firm pays dividend in the current year and zero otherwise; SPI_t [8] is special items scaled by lagged total asset.

To distinguish loss reversal probability between diversified firms and focused firms, we include variable $DIVER$ which represents a firm's diversity level. We use three ways to define $DIVER$. First, we use a dummy variable $DIVERSITY\_D$ equal to one if the firm has more than one business segment and zero otherwise. The dummy variable captures the aggregation difference between one-segment firms and multiple-segment firms. The second and third measurement focuses on the level of firms' segments diversity level and we are using

[7] Firms that pay dividends tend to be more profitable than firms not paying dividends (Fama and French 1999)

[8] We include the special item which is used in Li's (2011) forecasting model since a special item is transitory which results in losses being likewise transitory.

$HERF\_ASSET$ and $HERF\_SALES$ which are Herfindahl index by asset and sales. The Herfindahl index is calculated across N segment for each firm j as the sum of the squares of each segment i's total asset (sales) as a proportion of the firm level total asset (sales):

$$Herfindahl\ Index, H = \Sigma_{i=1}^{Njt}\left(\frac{x_{ijt}}{\Sigma_{i=1}^{Njt}X_{ijt}}^{2}\right) \quad (2)$$

where x_{ijt} is the segment level asset (sales) and X_{ijt} is the firm level total assets (sales). The asset (sales) based Herfindahl index reflects the degree to which total assets (sales) are diversified across firm's business segments with a range between zero to one. Focused firm has a Herfindahl index of one while the higher the firm diversifies, the lower the Herfindahl index is.

We construct our abandonment option following Lawrence et al., (2017). We define the firm's exercise abandonment option if both sales and the number of employees decrease compared to the previous period. We use the dummy variable equal to 1 if the firm exercises the abandonment option and 0 otherwise.

$$AB\_D_{t+1} = 1 \quad if\ Sales_{t+1} < Sales_t\ and\ Emp_{t+1} < Emp_t; 0\ otherwise. \quad (3)$$

To further control for firm difference between diversified firms and focused firms, we include several control variables: leverage ratio, Tobin's Q, investment and credit rating. First, leverage ratio ($LEVERGE_t$) is defined as total debt over total assets. Firms with high leverage ratio are likely to reduce their debt-financing costs by increasing the level of income-increasing accruals (Lim, Thong, and Ding 2008). Therefore, we control leverage ratio for any potential loss reversal due to earnings manipulation. Second, T_Q_t is defined as market value of firm divided by the book value of total assets. Focused firms tend to have better investment opportunity due to more growth options (Wernerfelt and Montgomery 1988, Lang and Stulz 1994, Berger and Ofek 1995, Servaes 1996, and Thomas 2002, Chen 2006). We use Tobin's Q to control for differences in inefficient investment caused by firm characteristics. Third,

$CAPX_t$ is defined as the capital expenditure scaled by total asset. Focused firms have higher investment size comparing to diversified firms, as is widely documented in the literature (Denis, Denis, and Sarin 1997, Chen 2006). Fourth, we control for a financial constraint effect by using credit rating CR_t from Compustat. Firms with higher credit rating are much more easily financed through cheaper external funds than are lower credit rating firms. Following Almeida et al. (2004) we use CR_t to define firms with a credit rating during sample periods as financially unconstrained firms while firms are financially constrained if they have no credit rating during the sample period.

Building on model (1), we add both dummy AB_D_{t+1} and interactive variable $AB\_DIVER_{t+1} = AB\_D_{t+1} * DIVER_t$ to capture the difference between diversified firms and focused firms in exercising abandonment option when suffering loss.

$$\begin{aligned} y_{t+1} = \alpha &+ \beta_1 EARN_t + \beta_2 PAST\_EARN_t + \beta_3 SIZE_t + \beta_4 SALESG_t + \beta_5 FIRSTLOSS_t \\ &+ \beta_6 LOSS\_SEQ_t + \beta_7 DIVDUM_t + \beta_8 SPI_t + \beta_9 DIVER_t + \beta_{10} LEVERGE_t \\ &+ \beta_{11} CAPX_t + \beta_{12} T\_Q_t + \beta_{13} CR_t + \beta_{14} AB\_D_{t+1} + \beta_{15} AB\_DIVER_{t+1} \\ &+ \varepsilon_{t+1} \qquad (4) \end{aligned}$$

4.5. Results

We first present the prevalence of firms' reporting losses, followed by descriptive statistics. In our main analysis, we show the relation between firm diversification and loss reversal profitability using the full sample. Then we introduce our abandonment option analysis in explaining the loss reversal difference in firm diversification. We next investigate the condition when firms are likely to exercise their abandonment option by dividing our sample into groups, according to whether the firm is suffering over-investment due to the higher agency problem of diversified firms comparing to the focused firms. Finally, we employ four econometric methods to control for endogeneity.

4.5.1 The prevalence of reporting loss by firm types

Figure 4.1 plots the percentage of U.S. firms from COMPUSTAT reporting annual losses from 1979 to 2016 by firm diversification. Overall, focused firms have a higher percentage of firms reporting losses than for the diversified firms. Both types of firm show similar trends of reporting losses through the full period. There is an increase of firms reporting losses from 9% for focused firms and 17% for diversified firms to the peak of 58% for focused firms and 38% for diversified firms in 2002. The percentage of loss reporting for both types of firms decreases until the reversion of the trend in 2007. This is consistent with the evidence presented by Li (2011) for firms overall. The percentage of loss reporting peaks in 2009. The two sharp increases of reporting losses in 2002 and 2009 are consistent with financial crises in those two years.

Figure 4.1

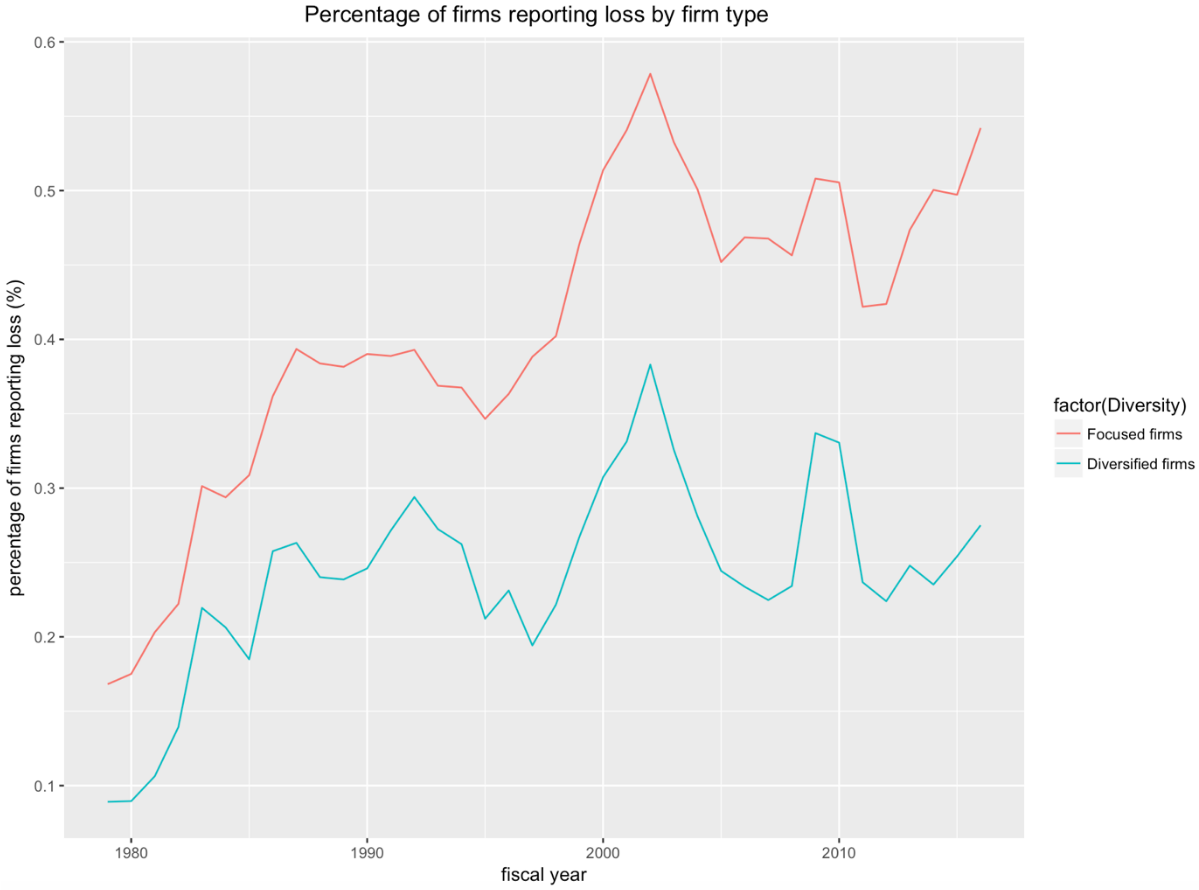

Figure 4.1 is based on data collected from Compustat over a period from 1979 to 2016. By distinguishing the firm's diversification, the figure demonstrates the percentage of firms reporting losses in each year. We define losses as income before extraordinary items and discontinued operations (annual Compustat data item #18). Firms with more than one segments are defined as diversified firms while focused firms are firms with only one segment.

Loss reversal is when firms become profitable after a previous year's losses. Figure 4.2 plots the percentage of loss reversal occurrence by firm's diversification. The percentage of loss reversal occurrence by each firm type is calculated from the number of firms that are loss-making in the current year and become profitable in the next year, divided by the total number of loss making firms in the current year. Both types of firm show a downward trend from 1979 to 2016. Overall, diversified firms have higher percentage of loss reversal occurrence than focused firms.

Figure 4.2

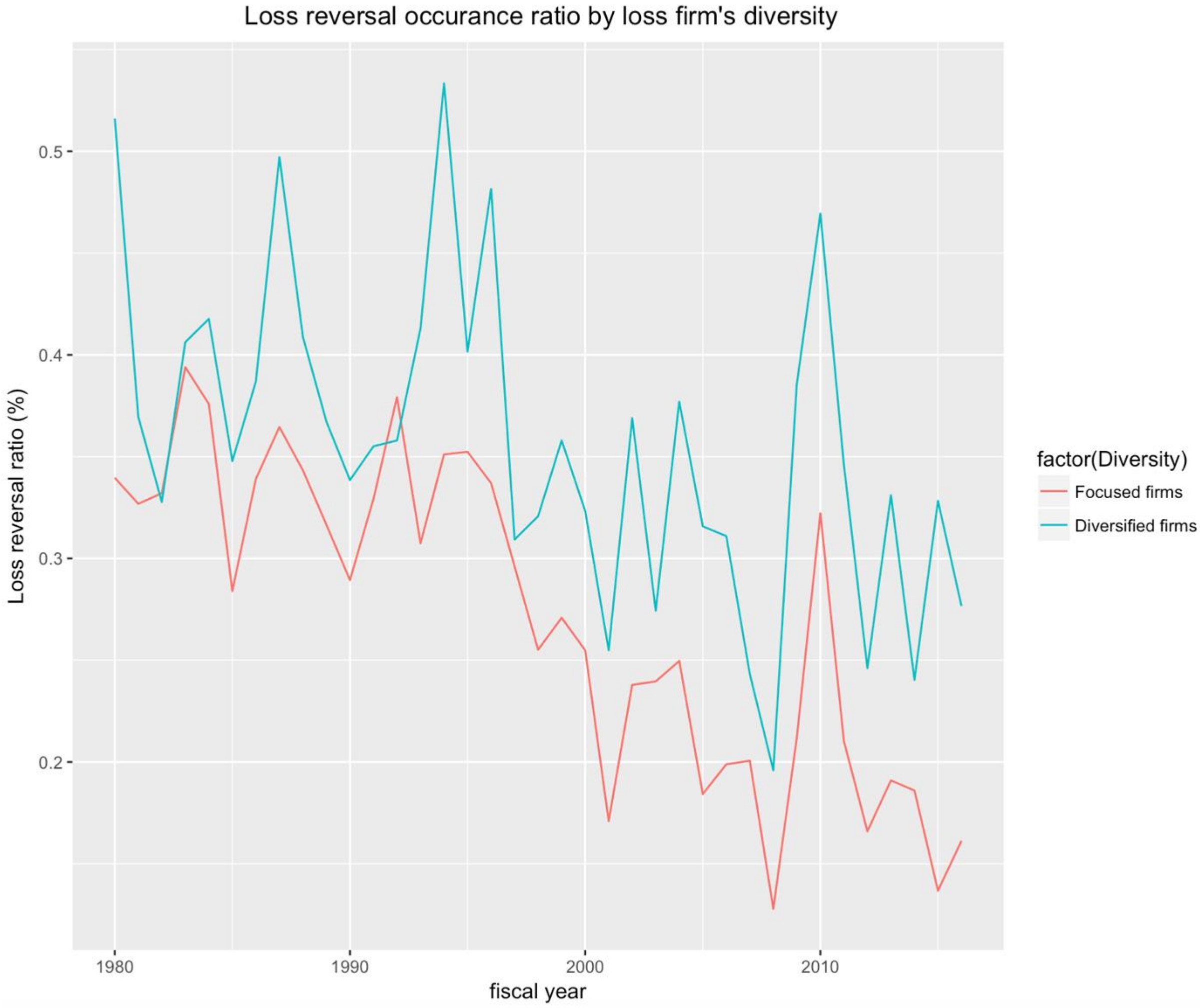

Figure 4.2 is based on data collected from Compustat over the period 1979 to 2016. By distinguishing the firm's structure, the figure demonstrates the loss reversal occurrence in each year. We define losses as income before extraordinary items and discontinued operations (annual Compustat data item #18). Firms with more than one segments are defined as diversified firms while focused firms are firms with only one segment. Loss reversal is defined when firms become profitable after a previous year's losses.

4.5.2 Descriptive statistics

Table 4.2 Panel A reports loss observations by firm type. We have 29290 focused firm-years and 9254 diversified firm-years. We summarize the control variables statistical description used in our main regression as follows. First, diversified firms have median (mean) size of 4.16 (4) in logarithm which is higher than the size of focused firms, whose median (mean) is 3.53 (3.43). Tobin's Q for focused firm is two times as high as for diversified firms. Capital expenditure to total asset are similar for diversified firms and focused firms. Diversified firms have higher leverage than focused firms (0.54 vs. 0.4) which is consistent with the prior finding that diversified firms have higher debt capacity due to the uncorrelated cash flows from different segments. Credit rating for diversified firm is better than for focused firms. Thus, it is necessary for us to include those variables to control any difference between diversified firms and focused firms.

Table 4.2

Data descriptive by the firm type

	Focused firm				Diversified firm			
	Obs.	Mean	SD	Median	Obs.	Mean	SD	Median
ROA	29290	-0.238	0.370	-0.118	9254	-0.125	0.237	-0.055
PAST_ROA	29290	-0.179	0.386	-0.046	9254	-0.049	0.232	0.008
SIZE	29290	3.533	1.992	3.429	9254	4.164	2.173	4.029
SALESG	29290	0.101	0.686	-0.015	9254	0.062	0.536	-0.014
FIRSTLOSS	29290	0.311	0.463	0.000	9254	0.450	0.498	0.000
LOSS_SEQ	29290	2.224	2.045	2.000	9254	1.431	1.757	1.000
DIVDUM	29290	0.224	0.417	0.000	9254	0.399	0.490	0.000
SPECIAL_ITEM	29290	-0.033	0.074	0.000	9254	-0.038	0.071	-0.008
LEVERAGE	29290	0.403	0.274	0.372	9254	0.540	0.244	0.566
CAPX	29290	0.055	0.084	0.027	9254	0.056	0.077	0.033
CREDIT	29290	0.393	0.488	0.000	9254	0.515	0.500	1.000
TOBIN'S Q	29290	1.813	3.648	0.675	9254	0.871	2.142	0.388
AB_D	29290	0.398	0.490	0.000	9254	0.465	0.499	0.000

This table summarizes the variables data descriptive used in our regression analysis. Variable definitions are provided in the notes to Table 4.1. Diversified firm is defined as the firms with multiple business segments while focused firms operate only one business.

4.5.3 Loss reversal and firm's structure

Table 4.3 reports the results of the logit regression model of Eq.1. Coefficients of most predictors inherited from Joos and Plesko (2005) and Li (2011) are consistent with their findings. Firms with higher current profitability, firms with shorter loss sequence, dividend-paying firms and firms with lower special items tend to have a higher probability of having a loss reversed in the next year. After introducing firm diversity related variables, the coefficient of firm size, first loss dummy variable and sales growth become insignificant. Loss sequences are negatively significant confirming that firms suffering shorter periods of loss are more likely to have losses reversed than for longer periods. The coefficients of the dividend payment dummy variable and the special items are all significant, consistent with Joos and Plesko (2005). Our new control variable credit rating is also positively significant, indicating that the higher the credit rating the higher the possibility to gain external funding to finance positive net present value projects and, as a result, higher possibility of achieving profits in the next period.

Our key variables are the firm's diversity related variables. Regression results (1) to (3) include the three versions of firm diversity. In equation (1), we include dummy variable equal to 1 if firms have more than one business segment and 0 otherwise. The coefficient is not significant, suggesting that there is no difference in the probability of reversal between diversified firms and focused firms. The marginal effects of firm diversification on the probability of loss reversal is 1.2% and significant at 90% level. One explanation of the insignificant coefficients of the firm diversification dummy variable is that firms with few segments (e.g. two segments) or multiple segments but running similar business lines (related diversifications) are close to the focused firms. Simply using a dummy variable to capture the aggregate difference may not be effective.[9] Therefore, we use a Herfindahl Index to capture the firm's diversity through firm asset and sales in equation (2) and (3). This index can reflect the correlation between business segments of the diversified firms. The index has a range from 0 to 1 and the higher the index, the lower the firm's diversity level. Particularly, the index value equals 1 for focused firms. As expected from Hypothesis 1, we find negative and significant coefficients of both Herfindahl Index variables suggesting that the higher the degree of firm's diversification, the higher probability of having a loss reversed in the following year. The marginal effect shows a consistent sign to further confirm the relation. The results suggest that the loss reversal difference between diversified firms and focused firms is not simply a fact of the difference in firm structure, but rather a matter of the diversity level of the firms' business segments. Building on these findings, we proceed to our main analysis of how firm structure affects the efficiency of exercising the abandonment option in the next section.

[9] We therefore use an alternative measure of diversification dummy which equals 1 if all the segments are operating in the same industry and 0 otherwise. We find this dummy variable is positive and significant which suggests that diversified firms have a higher probability of loss reversal than focused firms. As this alternative measure of diversity is consistent with the Herfindahl Index we used in the Regression (2) and (3), we drop this measure in the remaining tests.

Table 4.3

Loss reversal probability and firm structure: logistic model

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.247	0.072	0.001	***	-0.043	0.103	0.675		-0.006	0.104	0.956	
ROA	4.516	0.251	0.000	***	4.502	0.250	0.000	***	4.501	0.250	0.000	***
PAST_ROA	0.162	0.093	0.081	*	0.162	0.093	0.082	*	0.162	0.093	0.083	*
SIZE	-0.003	0.008	0.701		-0.006	0.008	0.510		-0.006	0.008	0.461	
SALESGROWTH	0.042	0.030	0.170		0.042	0.030	0.170		0.042	0.030	0.169	
FIRSTLOSS	-0.005	0.046	0.909		-0.006	0.046	0.899		-0.006	0.046	0.897	
LOSS_SEQ	-0.140	0.015	0.000	***	-0.139	0.015	0.000	***	-0.139	0.015	0.000	***
DIVDUM	0.159	0.035	0.000	***	0.151	0.035	0.000	***	0.150	0.035	0.000	***
LEVERAGE	-0.092	0.068	0.177		-0.110	0.068	0.106		-0.113	0.068	0.097	*
T_Q	0.013	0.009	0.161		0.013	0.009	0.157		0.013	0.009	0.151	
CAPX	-0.323	0.204	0.114		-0.302	0.204	0.139		-0.302	0.204	0.139	
SPECIAL_ITEM	-5.654	0.300	0.000	***	-5.646	0.300	0.000	***	-5.646	0.300	0.000	***
CREDIT_RATING	0.205	0.031	0.000	***	0.203	0.031	0.000	***	0.202	0.031	0.000	
DIVERSITY_D	0.033	0.035	0.347									
HERF_ASSET					-0.198	0.074	0.008	***				
HERF_SALES									-0.229	0.075	0.002	***
Diversity marginal effect	0.012	0.006	0.055	*	-0.05	0.013	0.000	***	-0.052	0.014	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 39362. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.5.4 Loss reversal and abandonment option

Table 4.4 presents the main analysis of our loss reversal model using the abandonment option. We follow Lawrence, Sloan and Sun (2017) in their measure of abandonment option dummy variable which equals 1 if both sales and number of employees decrease compared to the previous year. In Panel A, we present a descriptive statistical analysis to show the difference in frequency of exercising the abonnement option partitioned by firm structure. Consistent with our expectation in Hypothesis 2.1, we find that diversified firms have a significantly higher frequency of exercising the abandonment option than focused firms. For example, 39% of loss-making diversified firms have a positive abandonment option indicator compared with 33% of loss-making focused firms.

To compare the efficiency of exercising the abandonment option between diversified firms and focused firms (Hypothesis 2.2), we generate the interaction term of diversity variable and the abandonment option variable. In Table 4.4 Panel B, the results from regression (1) show significantly positive coefficients (0.23) of the interaction term, which suggests that diversified firms have higher probability of loss reversal using the abandonment option compared to the focused firms. The marginal effect noted in the table refers to the marginal effect of firm diversification when they exercise the abandonment option (AB_D=1). Holding other variables at their mean values, the predicted probability of loss reversal is 3.8% higher for diversified firms than focused firms when they both exercise the abandonment options. Similarly, using the Herfindahl diversity index in regression (2) and (3) we find consistent results (significant and negative coefficients of the interaction terms of -0.56 and -0.54) which means the higher the level of diversity the higher possibility that the firms to gain loss reversal through exercising abandonment options. The marginal effect at means of the Herfindahl index based interaction term is round -10% also support our results.

Table 4.4
Loss reversal and firm structure

Panel A: Descriptive Statistics for abandonment option variables by the firm structure

	Diversified firms			Focused firms								
Variable	*N*	Mean	Median	*N*	Mean	Median	Diff. in mean	p-value		Diff. in median	p-value	
EMP_DEC	9254	0.609	1.000	29290	0.528	1.000	0.081	0.000	***	0.000	0.000	***
SALE_DEC	9254	0.485	0.000	29290	0.462	0.000	0.024	0.000	***	0.000	0.000	***
AB_D	9254	0.390	0.000	29290	0.333	0.000	0.057	0.000	***	0.000	0.000	***

This table presents the statistical differences in means and medians for the abandonment options among loss firms by firm structure. Firms with more than one segment is defined as diversified firms and focused firms otherwise. EMP_DEC is 1 if the number of employees is lower than the previous year and 0 otherwise. SALE_DEC is 1 if sale is lower than previous year and 0 otherwise. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. We use two-sided t-tests to obtain the differences in means between diversified and focused firms. We use Wilcoxon signed-rank test to compare the median differences between diversified firms and focused firms. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.4
Loss reversal and firm structure

Panel B: Full logistic regression model

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.063	0.075	0.405		0.175	0.117	0.135		0.210	0.117	0.072	*
ROA	4.330	0.253	0.000	***	4.316	0.252	0.000	***	4.316	0.252	0.000	***
PAST_ROA	0.229	0.096	0.017	**	0.229	0.096	0.018	**	0.228	0.096	0.018	**
SIZE	-0.016	0.009	0.059	*	-0.019	0.009	0.027	**	-0.020	0.009	0.024	**
SALESGROWTH	-0.002	0.033	0.942		-0.002	0.033	0.943		-0.002	0.033	0.944	
FIRSTLOSS	-0.014	0.047	0.761		-0.016	0.047	0.737		-0.016	0.047	0.736	
LOSS_SEQ	-0.157	0.015	0.000	***	-0.156	0.015	0.000	***	-0.156	0.015	0.000	***
DIVDUM	0.169	0.036	0.000	***	0.160	0.036	0.000	***	0.159	0.036	0.000	***
LEVERAGE	0.192	0.072	0.008	***	0.173	0.072	0.016	**	0.172	0.072	0.017	**
T_Q	0.003	0.010	0.771		0.003	0.010	0.765		0.003	0.010	0.750	
CAPX	-0.785	0.214	0.000	***	-0.761	0.213	0.000	***	-0.761	0.213	0.000	***
SPECIAL_ITEM	-6.314	0.315	0.000	***	-6.313	0.315	0.000	***	-6.313	0.315	0.000	***
CREDIT_RATING	0.232	0.032	0.000	***	0.230	0.032	0.000	***	0.229	0.032	0.000	***
DIVERSITY_D	-0.009	0.044	0.845									
HERF_ASSET					-0.097	0.093	0.297					
HERF_SALES									-0.127	0.093	0.173	
AB_D	-1.235	0.043	0.000	***	-0.678	0.139	0.000	***	-0.697	0.140	0.000	***
AB_D*DIVERSITY_D	0.234	0.075	0.002	***								
AB_D*HERF_ASSET					-0.564	0.156	0.000	***				
AB_D*HERF_SALES									-0.540	0.157	0.001	***
Diversified &AB_D=1 marginal effect	0.038	0.009	0.000	***	-0.105	0.017	0.000	***	-0.100	0.017	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 38543. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.5.5 Over-investment and abandonment option

Diversified firms suffer more serious agency problems than focused firms. Managers may transfer resources from the profitable segments to the under-performing segments rather than abandoning those segments. We include a measure of the agency problem on the over-investment in diversified firms. In the situation of over-investment, we predict a lower difference or no difference in loss reversal probability through the abandonment option between diversified firms and focused firms (Hypothesis 3). We use the level of Tobin's Q as a proxy for over-investment. If a firm's Tobin's Q is lower than the median Tobin's Q in the same industry, this firm's asset generates less value than the industry which means the firms may engage in over-investments. We divide our sample into two groups by whether the firm's Tobin's Q is above or below the median industry Tobin's Q.

We estimate our regression by two groups and the results are reported in Table 4.5. For the firms engaging over-investment, we find that the coefficients of diversity-abandonment option interactive term (regression results (1), (2) and (3) in Table 4.5 A) are all insignificant. Although the marginal effects for Herfindahl Index diversity variable is significant at 95% level, both effects are much smaller than the marginal effect in the group with no over-investment problems (Table4.5B). These results suggest that when firms are engaging in over-investment, diversified firms cannot effectively liquidate their unprofitable assets to have losses reversed when compared to focused firms. In other words, when firms are suffering from agency problems, there is no difference in the firms' structure in terms of the efficiency of exercising the abandonment options since managers from either diversified firms or focused firms are reluctant to exercise the abandonment options.

In Table 4.5 B, we find that, in the group of firms not suffering from the over-investment problem, the coefficient and marginal effect of the interaction term of diversity and abandonment option is significantly positive in the diversity dummy version regression in regression (1) and negatively significant in regression (2) and (3). These results confirm the higher probability of an effective abandonment option for diversified firms comparing to focused firms.

Table 4.5

Loss reversal and firm structure: logistic model and agency problem

Panel A: Loss reversal and firm structure under over-investment

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.002	0.130	0.990		-0.014	0.178	0.938		0.055	0.177	0.758	
ROA	5.476	0.394	0.000	***	5.447	0.394	0.000	***	5.440	0.393	0.000	***
PAST_ROA	0.407	0.163	0.013	**	0.403	0.163	0.013	**	0.401	0.163	0.014	**
SIZE	-0.042	0.013	0.001	***	-0.046	0.013	0.000	***	-0.046	0.013	0.000	***
SALESGROWTH	-0.038	0.055	0.494		-0.039	0.055	0.485		-0.039	0.055	0.482	
FIRSTLOSS	0.030	0.066	0.644		0.028	0.066	0.669		0.028	0.066	0.670	
LOSS_SEQ	-0.145	0.021	0.000	***	-0.145	0.021	0.000	***	-0.145	0.021	0.000	***
DIVDUM	0.137	0.052	0.008	***	0.129	0.052	0.012	**	0.127	0.052	0.014	**
LEVERAGE	0.373	0.134	0.005	***	0.357	0.134	0.008	***	0.354	0.134	0.008	***
T_Q	0.241	0.087	0.006	***	0.244	0.087	0.005	***	0.246	0.087	0.005	***
CAPX	-1.504	0.428	0.000	***	-1.472	0.427	0.001	***	-1.467	0.426	0.001	***
SPECIAL_ITEM	-7.230	0.487	0.000	***	-7.220	0.487	0.000	***	-7.218	0.486	0.000	***
CREDIT_RATING	0.218	0.045	0.000	***	0.217	0.045	0.000	***	0.216	0.045	0.000	***
DIVERSITY_D	-0.070	0.062	0.257									
HERF_ASSET					0.024	0.131	0.855					
HERF_SALES									-0.044	0.131	0.738	
AB_D	-1.122	0.058	0.000	***	-0.743	0.184	0.000	***	-0.789	0.186	0.000	***
AB_D*DIVERSITY_D	0.148	0.101	0.142									
AB_D*HERF_ASSET					-0.389	0.207	0.061	*				
AB_D*HERF_SALES									-0.335	0.209	0.110	
Diversified & AB_D=1 marginal effects	0.012	0.012	0.37		-0.058	0.024	0.018	**	-0.051	0.025	0.044	**

This table presents the results of loss reversal probability with agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 18293. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.5

Loss reversal and firm structure: logistic model and agency problem

Panel B: Loss reversal and firm structure under no over-investment

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.063	0.103	0.541		0.139	0.167	0.405		0.151	0.168	0.368	
ROA	3.758	0.307	0.000	***	3.751	0.306	0.000	***	3.751	0.307	0.000	***
PAST_ROA	0.167	0.122	0.169		0.170	0.122	0.163		0.169	0.122	0.166	
SIZE	-0.005	0.012	0.672		-0.008	0.012	0.525		-0.008	0.012	0.511	
SALESGROWTH	-0.018	0.042	0.676		-0.017	0.042	0.684		-0.017	0.042	0.687	
FIRSTLOSS	-0.095	0.069	0.167		-0.096	0.069	0.163		-0.096	0.069	0.161	
LOSS_SEQ	-0.166	0.021	0.000	***	-0.165	0.021	0.000	***	-0.166	0.021	0.000	***
DIVDUM	0.190	0.051	0.000	***	0.178	0.051	0.001	***	0.180	0.051	0.000	***
LEVERAGE	0.556	0.121	0.000	***	0.538	0.121	0.000	***	0.540	0.120	0.000	***
T_Q	-0.007	0.011	0.532		-0.007	0.011	0.537		-0.007	0.011	0.545	
CAPX	-0.716	0.254	0.005	***	-0.696	0.254	0.006	***	-0.698	0.254	0.006	***
SPECIAL_ITEM	-6.111	0.412	0.000	***	-6.108	0.412	0.000	***	-6.105	0.412	0.000	***
CREDIT_RATING	0.241	0.046	0.000	***	0.237	0.046	0.000	***	0.236	0.046	0.000	***
DIVERSITY_D	0.038	0.062	0.535									
HERF_ASSET					-0.193	0.135	0.153					
HERF_SALES									-0.197	0.136	0.147	
AB_D	-1.374	0.067	0.000	***	-0.519	0.217	0.017	**	-0.511	0.220	0.021	**
AB_D*DIVERSITY_D	0.381	0.116	0.001	***								
AB_D*HERF_ASSET					-0.854	0.243	0.000	***				
AB_D*HERF_SALES									-0.860	0.246	0.000	***
Diversified &AB_D=1												
Marginal effects	0.072	0.014	0.000	***	-0.191	0.008	0.000	***	-0.191	0.008	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. The number of observation is 20249. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.6. Endogeneity in firm diversification

Firm diversification can be an endogenous choice by a firm (Campa and Kedia 2002, Graham et al., 2002). In this section, we introduce three econometric methods to control for endogeneity in our regression tests which are propensity score matching method, Heckman (1979) two-stage estimation and two-way fixed effect model. Each econometric method solves the endogeneity problem from a different perspective.

4.6.1 Propensity-score matching

We use propensity score matching model to control for sample selection bias due to observable characteristics. As noted earlier in this thesis, Roberts and Whited (2012) comment that propensity score matching is the most commonly used methodology to address endogeneity concerns due to its simplicity. Based on the one-equation system, the key benefit of using the matched sample for the regression analysis is that it avoids specification of the function Subrahmanyam, Tang, and Wang (2014). While the disadvantage is the that the potential outcomes are taken as independent of the treatment assignment is untestable. Therefore, we use propensity score matching as a robust test to control for the difference of loss reversal probabilities accompanying firm diversification. The propensity score match method conducts two stages to estimate the average treatment effect for the treated group (ATT). We define the treatment indicator equals 1 firm diversified firms and 0 for focused firms. The outcome variable is the loss reversal dummy equal 1 if loss firms becomes profitable in the next year and 0 otherwise. In the first stage, we predict the probability of diversifying by using a logit regression, where the dependent variable takes value of 1 for diversified firms and 0 for focused firms. We use all the control variables that we use in our main regression as the independent variables in the logit regression. In the second step, based on the propensity scores obtained from the predicted probabilities in stage one, we construct the common support sample (matched sample) by using the nearest neighbourhood matching. This matching is a one-to-one matching with replacement so that one diversified firm is matched with a focused firm with the nearest propensity score. Table 4.6 Panel A presents the statistic difference between diversified firms and focused firms using the matched sample. Compared with the unmatched sample, the mean difference of most control variables between diversified firms and focused firms becomes much smaller and statistically insignificant. This indicates that our propensity score match process is quite successful. Both the abandonment option and the loss reversal dummy variables for diversified firm are significantly higher than focused firms under this matched sample which is consistent with our prior analyses. The further results of regression tests under propensity score matched sample are described in the following tables examining the interactive difference of firms exercising abandonment option by firm diversity.

Table 4.6

The effect of diversification on loss reversal: propensity score matching estimations

Panel A: matched sample description

	Diversified firms	Focused firms (Control firms)	Difference	
ROA	-0.114	-0.122	0.008	
PAST_ROA	-0.035	-0.042	0.007	*
SIZE	4.002	3.946	0.056	
SALESG	0.079	0.057	0.022	*
FIRSTLOSS	0.462	0.459	0.003	
LOSS_SEQ	1.341	1.378	-0.037	
DIVDUM	0.412	0.402	0.010	
SPECIAL_ITEM	-0.035	-0.035	0.000	
LEVERAGE	0.555	0.551	0.004	
CAPX	0.058	0.06	-0.002	**
CREDIT	0.517	0.498	0.019	*
AB_D	0.386	0.35	0.036	***
Loss reversal	0.385	0.36	0.025	***
TOBIN'S Q	0.822	0.925	-0.103	

This table reports the statistic difference between diversified firms and focused firms using the propensity matched sample to estimate the average treatment effect for the treated group. The treatment indicator equals 1 for diversified firms and 0 for focused firms. Difference is the mean difference between the diversified and focused (control) firms by using the paired T-tests. Detailed propensity score matched steps are described in the main content. Variable definitions are provided in the notes to Table 4.1. Diversified firm is defined as the firms with multiple business segments while focused firms operate only one business sectors. There are 17410 matches. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.6
Probit regression: first-stage of Heckman Test

Panel B: Probability of diversify

	Method 1				Method 2			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.163	0.071	0.022	**	0.505	0.141	0.000	***
ROA(T-1)					0.282	0.082	0.001	***
ROA(T-2)					0.147	0.063	0.020	**
PAST_ROA	0.139	0.053	0.009	***				
SIZE	0.095	0.005	0.000	***	0.056	0.017	0.001	***
SIZE(T-1)					-0.003	0.013	0.850	
SIZE(T-2)					0.044	0.014	0.002	***
SALES_GROWTH	0.044	0.016	0.006	***				
GDP					-1.953	1.137	0.086	*
GDP_GROWTH					-0.009	0.001	0.000	***
FRAC_IDUSTRY	0.300	0.054	0.000	***	0.068	0.054	0.206	
FRAC_DIVER	1.900	0.098	0.000	***	2.567	0.101	0.000	***
DIVIDUM	0.254	0.022	0.000	***				
LEVERAGE	0.941	0.046	0.000	***				
SPECIAL_ITEM	-0.377	0.151	0.012	**				
TOBIN'S Q	-0.004	0.005	0.420					
CAPX	-0.367	0.145	0.011	**	-0.211	0.158	0.181	
CAPX(T-1)					-0.087	0.084	0.300	
CAPX(T-2)					-0.098	0.072	0.172	
CREDIT_RATING	0.208	0.021	0.000	***				
S&P					0.130	0.042	0.002	***
MAJOR_EXCHANGE					-0.068	0.023	0.003	***
FIRST_LOSS	-0.015	0.031	0.625					
LOSS_SEQ	-0.048	0.009	0.000	***				
CONSTANT	-2.134	0.054	0.000	***	-0.791	0.093	0.000	***

This table presents the estimates of probability of diversify through a probit model using two sets of variables. To fully capture the probability of diversify, firm level, year level and economy level factors are involved in the model. Method 1 focuses on the loss forecasting model while method 2 follows Campa and Kedia (2002)'s selection of variables from their first-stage Heckman probit model. Year dummies were included and not reported. The number of observation is 34820. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.6.2 Heckman two-stage estimation

We use Heckman (1979) two-stage estimation tests to control for the self-selection problem. In the first stage, we use a probit regression to estimate the probability of a firm choose to diversify. The “treatment effect” is based on the difference between diversified firms and the matched focused firms with similar propensity scores of the “decision to diversify”. We use two sets of selection variables in the first stage of Heckman tests based on a probit regression.

We obtain the Inverse Mills ratio based on the first set of selection variables (Method 1) in our main content and report the results based on the second sets of selection variables (Method 2) in the Appendix 4.1 (Panel A, B and C)[10].

The first set of selection variables (Method1) builds on Berger and Ofek (1995) and Villalonga (2004). We include their control variables (size, profitability and capital expenditure) along with additional control variables: dividend, leverage ratio, Tobin's Q, credit rating for a richer specification. These variables are consistent with all the control variables that are used in our main regression analysis[11]. We also include the two instruments used in Campa and Kedia (2002): Fraction of Diversified Firms in the Industry is the fraction of all the firms in the industry that are diversified firms (PNDIV). Fraction of Industry Sales by Diversified Firms is the fraction of industry sales accounted for by diversified firms (PSDIV).

For the second sets of selection variables (Method 2) used in Inverse Mills Ratio, we follow Campa and Kedia (2002)'s choice. We include lagged 1 and 2 years' variables of ROA, capital expenditure and size. S&P is a dummy variable that takes the value 1 if the firm belongs to the S&P industrial index, and 0 otherwise. We create the dummy variable MAJOR_EXCHANGE that takes the value 1 when the firm is listed on NYSE, Nasdaq, or AMEX, and 0 otherwise. We also include year and economic factors. GDP is Gross Domestic Product. GDP Growth is the growth rate of Gross Domestic Product.

We report the results of the first stage probit regression in Table 4.6A using the first set of selection variables. In the second-stage, based on the estimated probability of firms' choice to diversity, we use the estimated coefficients to construct Inverse Mills Ratios as an additional control variable which include the treatment for a firms' self-selection into diversification.

4.6.3 Fixed effect model

[10] The results based on the Inverse Mills ratio using the second set of selection variables are consistent with the results based on the first set of variables.

[11] Similar methodology can be found in Hann et al., (2013) for choice of variables in their Heckman tests.

We introduce a two-way fixed effect model using fixed firm effects to control for unobservable firm characteristics and year to control for time effects which affect the diversification decision. The fixed effect model can reduce the interfirm variability as explained by Campa and Kedia (2002).

4.6.4 Results

We present our results for the loss reversal model using three econometric methods in Table 4.6 Panel B. Since these take up considerable space, we only report our results using a diversity dummy as a proxy of diversified or focused firms. In the unreported tables, we also use our two Herfindahl-based measures for the level of diversification, the results are consistent with our results using a diversity dummy variable. Regression (1) shows the results using a propensity matched sample. Regression (2) reports the second-stage of Heckman estimation and Regression (3) presents the results of the two-way fixed effect model. We find the interaction term of diversity dummy with abandonment option dummy is positive and significant in all three regression models. The results suggest that diversified firms exercise abandonment option more efficiently than focused firms after controlling for endogeneity problems.

Similarly, in Table 4.6 Panel C and Panel D, we report the results of loss reversal analysis after controlling endogeneity problems by splitting the sample as if firms have agency problem and no agency problem, respectively. In Panel C, we find that among the firms with agency problems, diversified firms are no longer significantly more efficient than focused firms in exercising abandonment options. The superiority of exercising abandonment options hold in the firms with no potential agency problems as results shown in Panel D.

Table 4.6
Loss reversal and firm structure: logistic model with three econometric methods

Panel C: Loss reversal and firm structure full sample

	(1)				(2)				(3)			
	Propensity score match				Heckman Estimation				Fixed effect model			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.288	0.125	0.021	**	0.265	0.200	0.186					
ROA	4.986	0.483	0.000	***	4.221	0.255	0.000	***	4.240	0.164	0.000	***
PAST_ROA	0.138	0.175	0.429		0.216	0.097	0.026	**	0.195	0.093	0.036	**
SIZE	0.001	0.012	0.953		-0.022	0.011	0.045	**	0.010	0.009	0.305	
SALESGROWTH	0.002	0.050	0.966		-0.006	0.033	0.845		0.013	0.033	0.690	
FIRSTLOSS	-0.008	0.065	0.902		-0.006	0.048	0.894		-0.003	0.048	0.944	
LOSS_SEQ	-0.164	0.023	0.000	***	-0.154	0.015	0.000	***	-0.147	0.015	0.000	***
DIVDUM	0.195	0.045	0.000	***	0.150	0.040	0.000	***	0.122	0.037	0.001	***
LEVERAGE	-0.231	0.114	0.043	**	0.141	0.102	0.168		0.151	0.073	0.039	**
T_Q	-0.018	0.028	0.525		0.004	0.010	0.677		-0.003	0.009	0.720	
CAPX	-1.118	0.306	0.000	***	-0.751	0.218	0.001	***	-0.923	0.211	0.000	***
SPECIAL_ITEM	-6.888	0.509	0.000	***	-6.104	0.316	0.000	***	-6.201	0.266	0.000	***
CREDIT_RATING	0.198	0.043	0.000	***	0.233	0.034	0.000	***	0.163	0.037	0.000	***
DIVERSITY_D	-0.003	0.052	0.953		-0.022	0.045	0.630		0.001	0.044	0.983	
AB_D	-1.171	0.068	0.000	***	-1.230	0.044	0.000	***	-1.218	0.044	0.000	***
AB_D*DIVERSITY_D	0.188	0.092	0.040	**	0.221	0.076	0.004	***	0.225	0.076	0.003	***
INVERSE MILLS RATIO					-0.107	0.092	0.245					
Diversified & AB_D=1												
Marginal effects	0.023	0.010	0.019	***	0.020	0.007	0.003	***	0.020	0.012	0.027	**
Number of observation		34820				18508				29789		

This table presents the results of loss reversal logistic regression using three econometric methods to control for endogeneity problem. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Three measures of controlling for endogeneity problems are used. In regression (1), propensity score matched firms are selected based on propensity scores estimated from the model of probability of diversifying using all the variables in loss reversal model. In regression (2), we present the second-stage of Heckman Test with inverse Mills Ratio obtained from Table 6A using Method1. In regression (3), we use fixed effect model by controlling firm and year effects with firm and year effects unreported. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.6

Loss reversal and firm structure: logistic model with three econometric methods

Panel D: Loss reversal and firm structure under over-investment problem (high agency problem)

	(1)				(2)				(3)			
	Propensity score match				Heckman Estimation				Fixed effect model			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.358	0.207	0.084		0.145	0.304	0.634					
ROA	6.508	0.642	0.000	***	5.425	0.396	0.000	***	5.465	0.291	0.000	***
PAST_ROA	0.060	0.270	0.824		0.385	0.164	0.019	**	0.343	0.163	0.035	**
SIZE	-0.026	0.016	0.114		-0.045	0.017	0.008	***	-0.022	0.014	0.106	
SALESGROWTH	0.009	0.075	0.902		-0.038	0.056	0.501		-0.005	0.057	0.936	
FIRSTLOSS	-0.030	0.086	0.732		0.018	0.067	0.792		0.032	0.067	0.632	
LOSS_SEQ	-0.171	0.030	0.000	***	-0.150	0.022	0.000	***	-0.141	0.022	0.000	***
DIVDUM	0.146	0.062	0.018	**	0.132	0.057	0.021	**	0.089	0.053	0.097	*
LEVERAGE	-0.131	0.207	0.525		0.359	0.172	0.036	**	0.322	0.145	0.027	**
T_Q	0.223	0.138	0.106		0.249	0.088	0.004	***	0.250	0.098	0.011	**
CAPX	-1.317	0.562	0.019	**	-1.613	0.438	0.000	***	-1.793	0.422	0.000	***
SPECIAL_ITEM	-8.071	0.725	0.000	***	-7.052	0.488	0.000	***	-7.143	0.420	0.000	***
CREDIT_RATING	0.132	0.058	0.024	**	0.219	0.047	0.000	***	0.175	0.052	0.001	***
DIVERSITY_D	-0.056	0.072	0.439		-0.084	0.063	0.187		-0.066	0.063	0.297	
AB_D	-1.070	0.088	0.000	***	-1.114	0.059	0.000	***	-1.106	0.059	0.000	***
AB_D*DIVERSITY_D	0.121	0.120	0.313		0.134	0.102	0.189		0.139	0.102	0.174	
INVERSE MILLS RATIO					-0.063	0.138	0.645					
Diversified & AB_D=1 Marginal effects	-0.003	0.013	0.829		0.0004	0.011	0.969		-0.005	0.021	0.794	
Number of observation		16506				10084				18763		

This table presents the results of loss reversal logistic regression using three econometric methods to control for endogeneity problem with the firms having agency problems. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Three measures of controlling for endogeneity problems are used. In regression (1), propensity score matched firms are selected based on propensity scores estimated from the model of probability of diversifying using all the variables in loss reversal model. In regression (2), we present the second-stage of Heckman Test with inverse Mills Ratio obtained from Table 6A using Method1. In regression (3), we use fixed effect model by controlling firm and year effects with firm and year effects unreported. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.6

Loss reversal and firm structure: logistic model with three econometric methods

Panel E: Loss reversal and firm structure under no over-investment problem (low agency problem)

	(1)				(2)				(3)			
	Propensity score match				Heckman Estimation				Fixed effect model			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.054	0.179	0.764		0.041	0.273	0.881					
ROA	3.785	0.628	0.000	***	3.640	0.308	0.000	***	3.625	0.195	0.000	***
PAST_ROA	0.263	0.229	0.251		0.175	0.123	0.156		0.175	0.115	0.129	
SIZE	0.015	0.018	0.387		-0.008	0.015	0.583		0.025	0.014	0.063	*
SALESGROWTH	-0.060	0.071	0.398		-0.024	0.042	0.573		-0.007	0.042	0.863	
FIRSTLOSS	-0.011	0.101	0.913		-0.068	0.070	0.332		-0.074	0.070	0.288	
LOSS_SEQ	-0.148	0.036	0.000	***	-0.159	0.022	0.000	***	-0.152	0.021	0.000	***
DIVDUM	0.231	0.068	0.001	***	0.177	0.058	0.002	***	0.157	0.053	0.003	***
LEVERAGE	0.262	0.188	0.163		0.588	0.159	0.000	***	0.510	0.121	0.000	***
T_Q	-0.026	0.033	0.424		-0.006	0.011	0.598		-0.010	0.010	0.323	
CAPX	-1.202	0.379	0.002	***	-0.676	0.260	0.009	***	-0.742	0.246	0.003	***
SPECIAL_ITEM	-6.434	0.665	0.000	***	-5.921	0.413	0.000	***	-6.004	0.358	0.000	***
CREDIT_RATING	0.286	0.066	0.000	***	0.247	0.048	0.000	***	0.140	0.054	0.010	***
DIVERSITY_D	0.027	0.075	0.721		0.032	0.064	0.618		0.051	0.063	0.415	
AB_D	-1.336	0.110	0.000	***	-1.374	0.068	0.000	***	-1.358	0.069	0.000	***
AB_D*DIVERSITY_D	0.335	0.145	0.021	**	0.371	0.119	0.002	***	0.383	0.119	0.001	***
INVERSE MILLS RATIO					-0.056	0.126	0.660					
Diversified & AB_D=1 marginal effects	0.057	0.015	0.000	***	0.039	0.009	0.000	***	0.030	0.021	0.014	**
Number of observation		18281				8424				11026		

This table presents the results of loss reversal logistic regression using three econometric methods to control for endogeneity problem with the firms having no potential agency problems. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Three measures of controlling for endogeneity problems are used. In regression (1), propensity score matched firms are selected based on propensity scores estimated from the model of probability of diversifying using all the variables in loss reversal model. In regression (2), we present the second-stage of Heckman Test with inverse Mills Ratio obtained from Table 6A using Method 1. In regression (3), we use fixed effect model by controlling firm and year effects with firm and year effects unreported. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.7. Robustness Tests

4.7.1 Alternative measure of agency problem

We introduce two alternative measures of agency problem proxy for robustness checking of our analyses which are institutional ownership and dividend payout ratio. First, several empirical studies discuss the important role of large shareholders such as institutional investors to decrease the agency problem by monitoring the firm (Shleifer and Vishny 1986, Coffee 1991, Crutchley, Jensen, Jahera, Raymond 1999). Institutional investors own large portion of stocks which can effectively obtain firm-specific information and influence the managers' action through various methods such as using voting powers (Demsetz 1983, Graves and Waddock 1990, Pozen 1994, Sundaramurthy, Rechner, and Wang 1996).

Therefore, we use institutional ownership as an alternative proxy for the agency problem. We use a consistent methodology to define and split the sample into two groups. If a firm's institutional ownership is below the industry median in a year, the firm is classified into the group with high agency problem, lacking external monitoring; and the converse for firms above the industry median.

We report our results in Table 4.7. In Panel A, for the firms with low institutional ownership we find insignificant coefficients for the diversity and abandonment option interaction terms in all three regression models. In Panel B, for the sample of high institutional ownership, we find significant coefficients for the interaction term of diversity and abandonment option for all three diversity measures. Consistent with results from our main analysis, these findings suggest that without external monitoring, the superiority of abandonment options option exercise for diversified firms over focused firms is dissipated.

The inclusion of agency problems to explain the dividend puzzle has a long history in the literature (Easterbrook 1984, Jensen 1986, Fluck 1998, Fluck 1999, Zwiebel 1996). La Porta, Lopez-de-Silanes, Shleifer, and Vishny (2000) introduce two competing agency models of dividends. One predicts that stronger minority shareholder rights are related to higher dividend payouts since the stronger rights empower minority shareholders to pressure insiders to disgorge cash. In the other, the relation is reversed and insiders pay dividends to establish a reputation for good treatment for weak investors. Either way, dividends can decrease agency problems.

Therefore, we use firms' payout ratios as another proxy for agency problems. Once again, we split the sample into two groups, defining those with high agency problems as firms whose payout is below the industry median during a year, and the rest as firms with low agency problems. In Appendix 4.2, we see the same broad outcomes and interpretations for Panel A and Panel B as we saw previously for Table 4.7. Thus, consistent with our prior findings, the efficiency of diversified firms exercising the abandonment option is dampened, due to agency problems, compared to the focused firms.

Table 4.7

Loss reversal and firm structure: logistic model and agency problem

Panel A: Loss reversal and firm structure under low-external monitory (high agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.174	0.152	0.250		-0.100	0.248	0.688		-0.039	0.248	0.876	
ROA	4.407	0.441	0.000	***	4.418	0.442	0.000	***	4.405	0.441	0.000	***
PAST_ROA	0.124	0.160	0.438		0.125	0.161	0.436		0.123	0.160	0.441	
SIZE	-0.015	0.022	0.491		-0.014	0.022	0.528		-0.015	0.022	0.484	
SALESGROWTH	-0.004	0.064	0.946		-0.005	0.064	0.943		-0.004	0.063	0.948	
FIRSTLOSS	0.044	0.096	0.645		0.046	0.096	0.635		0.044	0.096	0.644	
LOSS_SEQ	-0.165	0.029	0.000	***	-0.165	0.029	0.000	***	-0.165	0.029	0.000	***
DIVDUM	-0.031	0.077	0.687		-0.028	0.077	0.713		-0.033	0.077	0.673	
LEVERAGE	0.336	0.141	0.018	**	0.342	0.141	0.015	**	0.334	0.141	0.018	**
T_Q	0.022	0.018	0.203	***	0.022	0.018	0.202		0.022	0.018	0.203	
CAPX	-1.571	0.464	0.001	***	-1.588	0.464	0.001	***	-1.576	0.464	0.001	***
SPECIAL_ITEM	-5.952	0.628	0.000	***	-5.973	0.628	0.000	***	-5.950	0.627	0.000	***
CREDIT_RATING	0.149	0.066	0.023	**	0.149	0.066	0.023	**	0.149	0.066	0.023	**
DIVERSITY_D	-0.094	0.092	0.306									
HERF_ASSET					0.271	0.208	0.191					
HERF_SALES									0.213	0.207	0.303	
AB_D	-1.293	0.086	0.000	***	-0.993	0.336	0.003	***	-0.803	0.333	0.016	**
AB_D*DIVERSITY_D	0.186	0.163	0.254									
AB_D*HERF_ASSET					-0.279	0.367	0.447					
AB_D*HERF_SALES									-0.494	0.364	0.175	
Diversified & AB_D=1 marginal effects	0.002	0.014	0.356		-0.003	0.021	0.598		-0.03	0.020	0.436	

This table presents the results of loss reversal probability with agency problems using the annual estimation of logistic regression. Low external monitory is used as the proxy of high agency problem which is defined as if the firm's institution ownership is below the industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 9513. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7

Loss reversal and firm structure: logistic model and agency problem

Panel B: Loss reversal and firm structure under high external monitory (low agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.296	0.149	0.047	**	0.385	0.219	0.079	*	0.407	0.218	0.062	*
ROA	6.779	0.580	0.000	***	6.740	0.579	0.000	***	6.741	0.579	0.000	***
PAST_ROA	0.374	0.218	0.085	*	0.377	0.218	0.084	*	0.374	0.218	0.085	*
SIZE	0.023	0.018	0.209		0.016	0.018	0.365		0.017	0.018	0.356	
SALESGROWTH	-0.021	0.063	0.735		-0.021	0.063	0.738		-0.021	0.063	0.738	
FIRSTLOSS	-0.281	0.084	0.001	***	-0.282	0.085	0.001	***	-0.281	0.085	0.001	***
LOSS_SEQ	-0.198	0.030	0.000	***	-0.197	0.030	0.000	***	-0.197	0.030	0.000	***
DIVDUM	0.278	0.063	0.000	***	0.263	0.063	0.000	***	0.264	0.063	0.000	***
LEVERAGE	0.012	0.140	0.934		-0.030	0.140	0.832		-0.028	0.140	0.843	
T_Q	-0.001	0.025	0.972		0.000	0.025	0.994		0.000	0.025	0.995	
CAPX	-1.155	0.434	0.008	***	-1.110	0.433	0.010	**	-1.112	0.433	0.010	**
SPECIAL_ITEM	-8.279	0.607	0.000	***	-8.248	0.607	0.000	***	-8.255	0.606	0.000	***
CREDIT_RATING	0.341	0.057	0.000	***	0.334	0.057	0.000	***	0.336	0.057	0.000	***
DIVERSITY_D	-0.054	0.077	0.478									
HERF_ASSET					-0.063	0.159	0.694					
HERF_SALES									-0.093	0.158	0.556	
AB_D	-1.261	0.078	0.000	***	-0.598	0.221	0.007	***	-0.675	0.224	0.003	***
AB_D*DIVERSITY_D	0.277	0.126	0.028	**								
AB_D*HERF_ASSET					-0.675	0.253	0.008	***				
AB_D*HERF_SALES									-0.580	0.256	0.023	**
Diversified & AB_D=1 marginal effects	0.082	0.042	0.000	***	-0.178	0.023	0.000	***	-0.177	0.024	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. High external monitory is used as the proxy of low agency problem which is defined as if the firm's institution ownership is above the industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 11092. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.7.2 Alternative abandonment option measures

Our main measure for the abandonment option follows the method of Lawrence et al., (2017). As they note in their paper, there are limitations to the method. The biggest problem in using decrease in sales is that this decrease could also be caused by the improvement of efficiency of operations and sticky costs effects. To mitigate this problem, we introduce two alternative measures in place of the decrease in sales. The first uses decrease in total assets, the second the decrease in sales growth from a forecasting model.

The first alternative measure is constructed as follows:

$$AB\_D_{t+1} = 1 \quad if\ Total\ asset_{t+1} < Total\ asset_t\ and\ Emp_{t+1} < Emp_t; 0\ otherwise. \quad (5)$$

The decrease in total assets can directly reflect the firm level abandonment activities including either disposing the entire segments or a fraction of the assets[12]. The key findings from using this alternative are consistent with our primary measure of the abandonment option, that diversified firms have a significantly higher probability of loss reversal compared to focused firms (Table 4.8 Panel A). Similarly, the efficiency of using the abandonment option is reduced under the over-investment samples (Appendix 4.3 Panel A and B).

The second measure is quite different from our primary measures. Returning to profitability from loss is a maintained hypothesis of financial reporting, embodied in "the going-concern" assumption (Joos and Plesko 2005). We use the possibility of exercising the abandonment option rather than the change in the numbers that have already happened. This method proceeds in two stages.

[12] Note that a similar approach used by Lawrence et al., (2017) is the existence of discontinued operations (known as Compustat acronym *do*). The key limitation of using discontinued operations is that, although IFRS and US GAAP treat discontinued operations differently, both accounting standards require discontinued operations as an entire business segment not the scaling back of a business unit.

During the first stage, we perform a sales growth forecast using an industry-specific model. The strategic management literature (Myers and Majd 1990, Berger, Ofek and Swary 1996) models the abandonment option as an American put option. The option is exercised when a loss occurs and future performance is under expectation. Instead of benchmarking past performance, our sales growth forecasting model predicts the future sales condition. If the industry condition is expected to be bad, firms are likely to leave this unprofitable industry and enter a new profitable industry (Fama 2000). In the second stage, we construct the abandonment option dummy variable equal to one if the predicted sales growth is negative and the firm suffers a loss in the current year, zero otherwise. Restrictions on both the firms' future and current performance are most consistent with the situation when firms will exercise the abandonment option.

We use dummy variable AB_D as an indicator of the existence of curtailments but using sales growth forecasting ($PREDGSL$).

To construct $PREDGSL$, we use the following simple first-order autoregressive model estimated by OLS regression on an industry-specific basis using 4-digits SIC (Standard Industry Code) as industry definition:

$$GSL_{t+1} = \mu_{j,T} + \nu_{j,T} GSL_{i,t} + \epsilon_{i,t}, \qquad (6)$$

where GSL_{T+1} is the growth in sales of firm i in year t, $\epsilon_{i,t}$ is the error term, and $t = T - 5, \dots, T-1$. The model parameters $\mu_{j,T}$ and $\nu_{j,T}$ are indexed by industry j and year T to highlight that the estimation is done on an industry-specific basis and for each year T using the previous five years of data. We then use the calculated estimation coefficient to generate the $PREDGSL$ for each firm-year $(i, T+1)$. We use an industry-specific model because previous work by Fairfield, Ramnath, and Yohn (2009) finds that sales growth forecasts are more accurate using an industry-specific model over the standard pooling regression (economy-wide model).

Thus, we construct our dummy variable AB_D to measure the existence of the curtailment, which is set to one when firms make losses during the current year and the predicted sales growth ($PREDGSL$) in the next year is smaller than 0. Formally:

$$AB\_D_{t+1} = 1 \quad if\ PREDGSL_{t+1} < 0; 0\ otherwise. \quad (7)$$

If a firm suffers loss and the predicted sales growth is negative, this provides a situation in which this firm should exercise its option to abandon the loss-making asset in order to prevent the loss from continuing in the next year. The key findings of using this alternative of an abandonment option are consistent with our primary measure of abandonment option that the diversified firms have a significantly higher probability of loss reversal compared to the focused firms (Table 4.8 Panel B). In addition, we find consistent results that the efficiency of diversified firms is reduced in the agency problem sample (Appendix 4.4 Panel A and B).

4.7.3. Alternative loss reversal model

In the second specification of our loss reversal model, we decompose ROA into two components: cash flow and accrual. Sloan (1996) shows that accrual component of earnings reflects the lower persistence of earnings comparing to the cash flow components. We define cash flow (CFO) as cash flow from operations as the difference between net income (annual Compustat item 172) and accruals. We define accruals (ACCR) as the change in current asset (Compustat item 4) – change in cash (Compustat item 1)-change in current liabilities (Compustat item 5) + change debt in current liabilities (Compustat item 34) + depreciation and amortizations (Compustat 14), deflated by the lagged total assets. We average the past five years CFO and ACCR as two additional control variables to capture the long-term effect of cash flow from operations and accruals on loss reversal following Joos and Plesko (2005).

$$
\begin{aligned}
y_{t+1} = \alpha &+ \beta_1 CFO_t + \beta_2 ACCR_t + \beta_3 PAST\_CFO_t + \beta_4 PAST\_ACCR_t + \beta_5 SIZE_t \\
&+ \beta_6 SALESG_t + \beta_7 FIRSTLOSS_t + \beta_8 LOSS\_SEQ_t + \beta_9 DIVDUM_t + \beta_{10} SPI_t \\
&+ \beta_{11} DIVER_t + \beta_{12} LEVERGE_t + \beta_{13} CAPX_t + \beta_{14} T\_Q_t + \beta_{15} CR_t \\
&+ \beta_{16} AB\_D_{t+1} + \beta_{17} AB\_DIVER_{t+1} \\
&+ \varepsilon_{t+1} \qquad (8)
\end{aligned}
$$

The findings in our alternative model analyses is consistent with our main regression's results. First, in Appendix 4.5 Panel A, we see the loss reversal probability is positively related to the firm's diversity level. Second, results suggest that diversified firms are more efficient than focused firms in exercising abandonment options in Appendix 4.5 Panel B. Third, this advantage of liquidating the loss-making assets in diversified firms over focused firms is reduced among the sample with high agency problems (Appendix 4.5 Panel C and D).

Table 4.8

Loss reversal and firm structure

Panel A: Logistic regression model with the asset-based abandonment option

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.125	0.058	0.032	**	0.200	0.094	0.033	**	0.245	0.094	0.009	***
ROA	4.272	0.189	0.000	***	4.259	0.189	0.000	***	4.261	0.189	0.000	***
PAST_ROA	0.252	0.089	0.004	***	0.251	0.088	0.005	***	0.248	0.088	0.005	***
SIZE	-0.005	0.007	0.450		-0.008	0.007	0.248		-0.008	0.007	0.226	
SALESGROWTH	0.010	0.026	0.707		0.010	0.026	0.701		0.010	0.026	0.710	
FIRSTLOSS	-0.021	0.037	0.562		-0.022	0.037	0.555		-0.022	0.037	0.550	
LOSS_SEQ	-0.148	0.012	0.000	***	-0.147	0.012	0.000	***	-0.148	0.012	0.000	***
DIVDUM	0.185	0.028	0.000	***	0.176	0.028	0.000	***	0.175	0.028	0.000	***
LEVERAGE	0.287	0.057	0.000	***	0.269	0.057	0.000	***	0.269	0.057	0.000	***
T_Q	-0.007	0.008	0.392		-0.007	0.008	0.404		-0.006	0.008	0.426	
CAPX	-0.826	0.163	0.000	***	-0.805	0.163	0.000	***	-0.807	0.163	0.000	***
SPECIAL_ITEM	-6.303	0.242	0.000	***	-6.301	0.242	0.000	***	-6.303	0.242	0.000	***
CREDIT_RATING	0.245	0.025	0.000	***	0.242	0.025	0.000	***	0.241	0.025	0.000	***
DIVERSITY_D	-0.038	0.036	0.285									
HERF_ASSET					-0.075	0.077	0.334					
HERF_SALES									-0.123	0.077	0.113	
AB_D	-1.377	0.032	0.000		-0.800	0.103	0.000	***	-0.852	0.104	0.000	***
AB_D*DIVERSITY_D	0.295	0.056	0.000	***								
AB_D*HERF_ASSET					-0.564	0.115	0.000	***				
AB_D*HERF_SALES									-0.502	0.116	0.000	***
Diversified & AB_D=1 marginal effects	0.027	0.006	0.000	***	-0.065	0.011	0.000	***	-0.062	0.012	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both total assets and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 38332. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.8
Loss reversal and firm structure

Panel B: Logistic regression model with abandonment option using GSL forecasting model

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.187	0.056	0.001	***	-0.011	0.081	0.894		0.024	0.082	0.771	
ROA	4.047	0.231	0.000	***	4.040	0.230	0.000	***	4.041	0.231	0.000	***
PAST_ROA	0.120	0.104	0.248		0.120	0.104	0.249		0.119	0.104	0.253	
SIZE	0.004	0.009	0.667		0.002	0.009	0.831		0.002	0.009	0.869	
SALESGROWTH	0.014	0.030	0.633		0.015	0.030	0.630		0.015	0.030	0.627	
FIRSTLOSS	-0.101	0.051	0.047	**	-0.102	0.051	0.045	**	-0.102	0.051	0.044	**
LOSS_SEQ	-0.142	0.016	0.000	***	-0.141	0.016	0.000	***	-0.141	0.016	0.000	***
DIVDUM	0.178	0.038	0.000	***	0.168	0.038	0.000	***	0.169	0.038	0.000	***
LEVERAGE	0.275	0.087	0.002	***	0.255	0.087	0.003	***	0.256	0.087	0.003	***
T_Q	0.000	0.008	0.965		0.000	0.008	0.967		0.001	0.008	0.948	
CAPX	-0.283	0.181	0.119		-0.260	0.182	0.152		-0.266	0.181	0.142	
SPECIAL_ITEM	-5.845	0.300	0.000	***	-5.841	0.300	0.000	***	-5.843	0.300	0.000	***
CREDIT_RATING	0.193	0.034	0.000	***	0.190	0.034	0.000	***	0.189	0.034	0.000	***
DIVERSITY_D	0.071	0.041	0.083	*								
HERF_ASSET					-0.295	0.088	0.001	***				
HERF_SALES									-0.306	0.089	0.001	***
AB_D	-0.155	0.081	0.056	*	0.519	0.280	0.064	*	0.593	0.277	0.032	**
AB_D*DIVERSITY_D	0.397	0.145	0.006	***								
AB_D*HERF_ASSET					-0.637	0.311	0.041	**				
AB_D*HERF_SALES									-0.724	0.308	0.019	**
Diversified & AB_D=1												
Marginal effect	0.030	0.007	0.002	***	-0.071	0.015	0.000	***	-0.070	0.015	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if model-predicted sales growth decreases comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 39257. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.8. Conclusion

We address the role of firm structure in explaining a firm's loss reversal probability using U.S. firms over the period 1980 to 2016. Compared to focused firms, firms with higher diversity level have higher probability of returning to profit after suffering losses.

Inspired by Hayn (1995) on the abandonment option for loss-making firms, we find that diversified firms have higher efficiency in exercising their abandonment options in order to reverse losses in the next period. Diversified firms have multiple business segments where cash flows are not perfectly correlated. After abandoning their loss-making assets or segments, they can still generate profit through other segments. Focused firms on the other hand are less likely to return to profit after liquidating their unprofitable assets because of their undiversified business lines.

However, diversified firms tend to have agency problems, leading to over-investment in unprofitable projects. The second part of our analysis provides clear evidence that over-investment impairs the efficiency of diversified firms in exercising their abandonment options. Due to the higher agency problems, managers from diversified firms tend to not abandon all the over-invested and loss-making projects (while engaged in empire building) making it more difficult for the firms to return to profit in the next period.

This chapter considers the real actions of firms dealing with loss. It extends the loss reversal model by including a firm's diversity as an important factor. More importantly, we explain the importance of firms' structures and show how their abandonment options are related to firm performance. In other words, given the irreversible nature of the abandonment option, we investigate how efficient firms can use this option to avoid losses in a going-concern strategic view. We show that a firm's diversification level is an important factor that determines this efficiency whereby diversified firms have higher flexibility in liquidating under-performing assets than focused firms.

More generally, we highlight the diversification discount literature by indicating that agency problems can be influential in the efficiency of firms' decisions to exercise their real options. The higher level of reluctance among managers for acting to liquidate in the diversified firms can be value-destroying, neutralizing their primary advantage of exercising the abandonment options.

Chapter 5

Firm diversification and dividend payout policy

5.1. Introduction

Study of firm diversification has a long history. Early studies focus on the relationships between diversification and the firm's performance. Evidence is found that product diversification cannot increase profitability (Grant, Jammine and Thomas 1988). Much empirical work investigates how firm structure differences affect the ways of distributing earnings, especially how firms make use of their retained earnings. Diversified firms tend to hold less cash (Opler et al., 1999, Subramaniam et al., 2011, Tong 2011), or over-invest (Berger et al., 1995). The effect of firm structure on dividend policy remains an interesting question.

The classic work by Modigliani and Miller (Modigliani and Miller 1958, Miller and Modigliani 1961) models a firms' dividend payout policy as having no influence on its capital structure or shareholders' wealth, with a constant investment policy in a frictionless world. Once we move way from this idealized perfect market and enter the real world of taxes, bankruptcy costs and other frictions, the question of how firms choose their dividend policy becomes open again and has been discussed in a voluminous literature.

One explanation of the dividend puzzle is the signalling hypothesis. Paying dividends can be used as a signal of the firm's future performance (Bhattacharya 1979, John and Williams 1985, Miller and Rock 1985, Ambarish, John, and Williams 1987). In general, dividend changes are positively associated with a firm's stock returns. However, studies show mixed results which do not support the relation between dividend changes and future profitability (DeAngelo, DeAngelo, and Skinner 1996, Benartzi, Michaely and Thaler 1997).

The use of agency problems to explain the dividend puzzle has become popular (Easterbrook 1984, Jesen 1986, Fluck 1998, Fluck 1999, Zwiebel 1996). La Porta, Lopez-de-Silanes, Shleifer, and Vishny (2000) introduce two competing agency models of dividends. On the one hand, the outcome model predicts stronger minority shareholder rights are related to higher dividend payout since the stronger rights empower minority shareholders to pressure insiders to distribute cash. On the other hand, the relation is reversed under the substitute model. Insiders pay dividends in order to establish a reputation for weak-investor protections.

We contribute to the dividend payout literature by exploring an un-researched area where a firm's structure can affect dividend policy. The contributions of our study to the literature are threefold. First, prior studies on firm diversification investigate how a firm's structure affects performance (Lang and Stultz 1994, Palich, Cardinal and Miller 2000), investment opportunity (Berger, et al 1995), and cash holding (Opler et al., 1999). Our study considers the dividend policy as an alternative channel whereby the firm is distributing its excess earnings.

Second, dividend policy is closely related to agency problems. Since diversified firms are more likely to have severe agency problems because managers from sub-divisions are involved in over-investment due to empire building (Jesen and Mackling 1976, Jensen 1986, Berger et al., 1995) and inefficient use of resources due to cross-subsidization (Shin and Stulz 1998, Rajan et al., 2000, Berger et al., 1995). Two widely-used competing models by Porta et al., (1995) predict the opposite relation between agency problem and dividend payout. Our research complements the agency literature on firm diversification by investigating the ratio of dividend payout.

Third, dividend payout serves a crucial role in the firm's financing strategy. Miller and Modigliani (1961) point out that investors interpret a dividend change as a signal of the change in manager's view of the firms' future performance. The dividend signaling model supposes that a reduction in dividend conveys the managerial pessimism about the firm's future (John and William 1985, Left and Zmijewski 1994, Jensen, Lundstrum and Miller 2010). The classic pecking order model of Myers (1984) points out there are financing costs concerns such as transactions costs and asymmetric information problem, especially for firms that are financed externally with risky debt or securities. Increased dividends can raise the transaction cost of external financing (Rozeff, 1982). Diversified firms have lower issue with financing shortage since segments have sufficient internal capital and positive NPV investment projects can be cross-financed with capitals within segments. In addition, internal capital markets can create firm value if firms are externally financial constrained. By investigating the financially constrained firms, our study supports the evidence of the efficient internal capital market for diversified firms in dividend policy.

To understand how dividend payout policies are related to firm structure, we use a sample of firm-years from 1980 to 2016. In general, we find that diversified firms have significantly higher dividend payout ratios than focused firms.

To find out why diversified firms pay higher dividends than focused firms, we test two hypotheses. First, firm diversification is associated with higher agency problems due to empire building (Jensen 1986) or over-investment (Berger et al., 1995). We divide our sample into groups by the level of agency problems using free cash flow and Tobin's Q as two alternative measures. We find that among the firms with high agency problems, diversified firms pay significantly higher dividends than focused firms which fits with the substitute hypothesis of Porta et al., (1995).

Second, we test the hypothesis that diversified firms pay higher dividends because of their large and efficient internal capital (Weston 1970, Stulz 1990, Stein 1997) especially when firms are financially constrained. We use coverage rate ratio and cash flow volatility as two proxies for financial constraint. Our results suggest that in the group of financially constraint firms, diversified firms pay higher dividends than focused firms, indicating that this excess dividend is generated through their large internal capital market.

Our regression models are controlled for firm and year effects. In the robustness tests, we use propensity score match and Heckman two-stage tests to control for endogeneity problems. We use two additional proxies for agency problem measures which are firm size and age. Large and mature firms are more likely to have fewer good investment opportunities than the small young firms. Therefore, managers in the large and mature firms are more likely to concoct agency problems. We find that among the large and mature firms, diversified firms have higher dividend payout ratios than focused firms, supporting the substitute hypothesis.

We review the relevant literature on firm diversification and dividend policy in Section 2. In Section 3, we present our two main hypotheses. We introduce our sample and explain the empirical model in Section 4. In Section 5 we assess our findings. We include robustness tests to support our findings in Section 6. In Section 7 we conclude.

5.2. Hypothesis design

Our primary goal in this chapter is to investigate the importance of firm structure in payout policy. If diversification plays a crucial role in a firm's performance as well as the managers' strategic decision, whether firms are diversified should have a significant impact on the firms' dividend policy. There should be at least two impacts of firm diversification on payout policy which are the agency problems and the internal capital market.

5.2.1 Agency problem hypothesis

Conflicts of interest between managers and shareholders can be at the centre of agency problems. Managers can use free cash flows to invest in unprofitable projects for personal benefit (Jensen and Meckling 1976) and divert corporate resources to themselves (Jensen 1986). Therefore, shareholders may insist that managers have the firm pay dividends to reduce discretionary cash under managers' control. This can force firms to seek external funding and become exposed by monitoring in the external market (Rozeff 1982, Esterbrook 1984), consistent with the substitute hypothesis referred by Parta et al., (2000).

Firm diversification is associated with increased agency problems, with segment managers tending to lobby and grab more firm-level resources to suit their self-interests (Rajan et al. 2000). Top management in the diversified firms might overinvest since they have higher opportunities to undertake projects and resources (Stulz 1990, Berger et al., 1995, Matsusaka and Nanda 1997). Consequently, agency problems destroy firms' value. Subranmaniam, Tang, Yue, and Zhou (2011) find that diversified firms hold less cash in order to mitigate the agency costs. It is called as the influence cost theory suggesting that diversified firms take effective actions to reduce cash holding which eliminates the fight over resources due to the empire building. If the diversified firms intend to hold less cash due to reducing the agency costs, we would also expect an increase of dividend payout ratio which is consistent with the substitute theory by La Porta, Lopez-de-Silanes, Shleifer and Vishny (2000). Substitute theory argues that managers payout dividend in order to establish repuations for good treatment of minority shareholders. To test this hypothesis, we use two agency proxies to dividend the sample into subgroups. First, we classify firms free cash flow. High free cash flow (if firms free cash flow is above the industry median) increase the marginal agency costs by providing the incentives for firms to use the excess resources to engage additional investments (more likely to over-investmnet). Second, we use Tobin's Q as an alternative measure of agency problem. Firms with low Tobin's Q (if firm's Tobin's Q is below the industry median) face less promising investment opportunities. Therefore, we expect that diversified firms eliminate cash holdings by distributing higher dividend payouts than focused firms in order to release the agency.

We proceed to our first hypothesis as follows:

H1: Firm diversification increases the level of dividend pay-out ratio to reduce the agency costs.

5.2.2 Internal capital market hypothesis

Compared to focused firms, diversified firms have a larger and more efficient internal capital market for transferring more resources to the divisions with higher investment opportunities (Li and Li 1996, Weston 1970, Stulz 1990, Stein 1997). This pooled internal capital markets suggests that diversification creates value since it can avoid the problem of foregone good investment opportunities due to the insufficient funds.

Chay and Suh (2009) argue that firms with high cash flow volatility are expected to rely more on internal funds and to pay low dividends due to costly external financing and the fear of future cash shortfall. We expect, therefore, that among financially constrained firms, diversified firms can fund more good investment projects through internal capital market and pay higher dividends than focused firms. This hypothesis is also consistent with the dividend signalling hypothesis. Firms pursue the stability of their stocks while firms can pay high dividends, to signal future strong performance and stabilize their stocks. On the other hand, among the financially unconstrained firms, either diversified firms or focused firms can more readily obtain external funds for their good investment projects. Therefore, dividends can be paid with sufficient financing internally and externally. We use two proxies, coverage ratio and cash flow volatility, to measure the firm’s level of internal and external financial constraint, respectively. We expect for either proxy, diversified firms can still pay higher dividend for the firms that are facing financial constraint issues.

Therefore, we proceed to our second hypothesis as follows:

H2: Firm diversification increase the level of dividend payout ratio in financially constrained firms through internal capital market.

5.3. Data and sample selection

We collect data from three sources via Wharton Research Data Services (WRDS). Firm level and business segment data are obtained from Compustat North America annual fundamentals file and segment files. Stock returns data used in portfolio constructions come from the Centre for Research in Security Prices (CRSP) monthly stock file. We sample for our model from 1979 to 2016. Firms were required to report segment information after December 15, 1977 by Statement of Financial Accounting Standard (SFAS) No.14.

We exclude financial firms (SIC from 600 to 6799). The segment data are merged with the firm level data to construct our matched samples. We remove any mismatched observations which have the summation of segment sales (segment asset) more than 101% or less than 99% (more than 125% or less than 75%) of the firm's total sales (total assets). A detailed description of our variables is provided in Table 5.1. For all the balance sheet data constructing accruals, we replace the missing values with zero following Ball et al. (2006). For all our empirical tests, we trim all continuous-value variables to the 1^{st} and 99^{th} percentiles.

TABLE 5.1
Variable definitions

Variable name	Description	Computation / WRDS mnemonic
(USD million)		
IB	Income before extraordinary items	IB
TA	Total assets	AT
NI	Net income	NI
DIV	Total dividend	DIV
SALES	Sales/Turnover (net)	SALE
DIV	Dividend payout ratio	Total dividend (DIV) scaled by net income (NI)
ROA	Return on asset	Income before extraordinary item (IB) scaled by Total assets (AT) lagged by one year
MV	Market value of firm	Price_Fiscal Year_Close (PRCC_F) * Common shares oustanding (CSHO)
SIZE	Firm size	Log of MV
LEVERAGE	Leverage level	[Long- term debt (DLTT) + Current liabilities (LCT)] / [Long-term debt (DLTT) + Current liabilities (LCT) + Market value (MV)]
T_Q	Tobin's Q	Market value (MV) / Total assets (TA) lagged by one year
CRD	Credit Rating	*Equal to one if firm has a credit rating, zero otherwise*
RD	Research and development expense	Research and development (RD) scaled by Total asset (AT) lagged by one year
CHE	Cash Holding	Cash holding (CHE) scaled by Total assets (AT) lagged by one year
RE	Retained earning	Retained earning (RE) scaled by Total assets (AT) lagged by one year
CFO volatility	Cash flow from operating activities	Standard deviation of CFOs from past five years
FCF	Free cash flow	Operating income before depreciation (OIBDP)- interest expense(XINT)- taxes(TXT-Change in TXDITC)- preferred dividends (DVP)- common dividends (DVC)
COVERAGE ratio	Coverage ratio	Net income (NI)/ interest expense (XINT)
HERF(TA)	Herfindahl index by total asset	
HERF(SALE)	Herfindahl index by sales	
Diver_D	Diversity dummy variable	*Equal to one if firm has more than one business segments, zero otherwise*

† If the data items for preferred stock, long-term debt, debt in current liabilities, minority interest and cash and short-term investments are not available, they are assumed to equal zero.
‡ If the data items from balance sheet accounts are not available, they are assumed to equal zero.

5.4. Empirical model

Our empirical model investigates the relation between dividend payout ratio and firm structure by including all the necessary firm level control variables as follows:

$$DIV_t = \alpha + \beta_1 EARN_t + \beta_2 SIZE_t + \beta_3 SALESG_t + \beta_4 DIVER_t + \beta_5 LEVERGE_t + \beta_6 CAPX_t + \beta_7 T\_Q_t + \beta_8 CRED\_R_t + \beta_9 CASH_t + \beta_{10} RE_t + \beta_{11} CFO_t + \beta_{12} R\&D_t + \varepsilon_{t+1}$$

where our dependent variable DIV_t is the dividend payout ratio which is, the total dividend divided by net income. $EARN_t$is the firm's profitability measurement using return on asset (ROA). This is defined as annual income before extraordinary items and discontinued operations (annual Compustat item #18) deflated by lagged total asset (annual Compustat item #6). $SIZE_t$ is the logarithm of market value of equity (annual Compustat data item#199*annual Compustat data item#25). $SALESG_t$ is the percentage growth in sales over year *t*; $LEVERGE_t$ is defined as total debt over total asset. $CAPX_t$ is defined as the capital expenditure scaled by lagged total asset. T_Q_t is defined as market value of firm divided by the book value of total assets. We include two dummy variables: $R\&D_t$ is equal to one if the firms have research and development expense and zero otherwise. $CRED\_R_t$ is equal to one if firms have a credit rating and zero otherwise. We also include two additional important variables to control dividend payments which are cash holding ($CASH_t$), retained earnings (RE_t), and cash flow from operating activities (CFO_t) all scaled by lagged total assets.

To distinguish the dividend payout ratio between diversified firms and focused firms, we include variable $DIVER_t$ which represents the firm's diversity level. We use three ways to define $DIVER_t$. First, we use a dummy variable equal to one if the firm has more than one business segment and zero otherwise. The dummy variable captures the aggregation difference between one-segment firms and multiple-segment firms. The second and third measurement focus on the level of firms' segments diversity level and we use Herfindal index (AT) and Herfindal index (Sales). The Herfindal index is calculated across *N* segment for each firm j as the sum of the squares of each segment i's total asset (sales) as a proportion of the firm level total asset (sales):

$$Herfindahl\ Index, H = \Sigma_{i=1}^{Njt}(\frac{x_{ijt}}{\Sigma_{i=1}^{Njt}X_{ijt}}^{2}) \quad (2)$$

where x_{ijt} is the segment level asset (sales) and X_{ijt} is the firm level total assets (sales). The asset (sales) based Herfindahl index reflects the degree to which total assets (sales) are diversified across firm's business segments with a range between zero to one. Focused firm has a Herfindahl index of one while the higher the firm diversifies, the lower the Herfindahl index is.

5.5. Empirical results

5.5.1 Descriptive statistics

Table 5.2 presents summary statistics for key variables used in our model and the difference between the group of firms that are diversified and the group of focused firms. First, diversified firms have mean dividend payment of 22% of their net income which is much higher than the dividend payout ratio for the focused firms (mean 10%). In addition, diversified firms and focused firms differ in possible determinants of dividend payout ratio. Diversified firms have lower cash holdings than focused firms (mean cash holding of 8% of their asset for diversified firms versus 19% for focused firms). This difference in cash holdings is consistent with previous studies (Opler 1999, Subramaniam, Tang, Yue and Zhou 2011, Tong 2011). Conversely, diversified firms hold higher retained earnings than focused firms (median of 22% for diversified firms comparing to 8%). We also find that statistically diversified firms and focused firms have similar investment with a median of 6% for diversified firms and 5% for focused firms.

Table 5.2

Data descriptive by firm type

	Focused firm (N= 57065)			Diversified firm (N= 26075)			Difference	
	Mean	SD	Median	Mean	SD	Median	Mean	
DIV PAYOUT	0.097	0.298	0.00	0.219	0.385	0.074	-0.122	***
ROA	-0.071	0.374	0.03	0.016	0.183	0.043	-0.087	***
SIZE	4.183	2.084	4.038	5.182	2.219	5.224	-0.999	***
SALESGROWTH	0.212	0.633	0.10	0.156	0.441	0.087	0.056	***
LEVERAGE	0.375	0.243	0.338	0.452	0.214	0.445	-0.077	***
CAPX	0.087	0.110	0.051	0.078	0.082	0.056	0.009	***
T_Q	1.790	3.033	0.900	0.988	1.524	0.629	0.802	***
CASH HOLDING	0.189	0.320	0.078	0.08	0.128	0.042	0.109	***
RETAINED EARNINGS	-0.705	2.956	0.08	0.013	1.358	0.215	-0.718	***
CFO	-0.159	0.426	-0.052	-0.057	0.216	-0.019	-0.102	***

This table summarizes the variables data descriptive used in our regression analysis including number of observations (Obs.), mean, standard deviation (SD) and median. Variable definitions are provided in the notes to Table 5.1. Diversified firm is defined as the firms with multiple business segments while focused firms operate only one business sectors. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

For firm level control variables, we find that diversified firms have a slightly higher leverage ratio (median of 0.4 versus 0.3) and size (median of 5.2 versus 4.0) while much lower Tobin's Q (median 0.6 versus 0.9) and sales growth (median of 9% versus 10%) than focused firms.

5.5.2 Regression evidence

In this section, we test whether diversified firms have higher dividend payout ratio than focused firms after controlling for firms' specific characteristics and the previously found determinants of dividend payments

5.5.2.1 Firms structure and dividend payout ratio

Table 5.3 presents the main regression analysis of the relationship between dividend payout ratio and diversification. We use the double fixed effect model to control for firm and year effects. We calculate standard errors by clustering firm and year. Table 5.3 Regression (1) uses a diversification dummy, which is one for diversified firms and zero for focused firms. This shows that diversified firms have a significantly higher dividend payout ratio than focused firms after controlling all the previously established determinants of dividend payout ratio such as cash holding, retained earnings, investment as well as the firm level variables such as leverage, Tobin's Q, size and credit rating.

For Table 5.3 Regression (2) and Regression (3), we use an alternative firm structure variable in place of a diversity dummy variable. We use a Herfindahl Index based on either assets and sales to determine the firms' level of diversify which means the higher the index, the lower the diversify level. Consistent with the findings in Regression (1), we find that firms with higher level of diversify have significantly higher level of dividend payout ratios.

From the three regressions' results of Table 5.3 we confirm that earnings, size and CFO have positive and significant effect on the dividend payout ratios. Firms with high leverage, high Tobin's Q and high credit rating pay lower dividend relative to their earnings which are all consistent with the previous studies.

Overall, we find that diversified firms pay higher dividends relative to their earnings than do focused firms. These findings are consistent with the substitute theory of dividend policy from the previous studies.

Table 5.3

Dividend payout ratio and firm structure: double fixed effect model

Full sample	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.031	0.005	0.000	***	0.031	0.005	0.000	***	0.031	0.005	0.000	***
SIZE	0.007	0.002	0.006	***	0.006	0.002	0.010	**	0.006	0.002	0.009	***
SALESGROWTH	-0.003	0.002	0.123		-0.003	0.002	0.118		-0.003	0.002	0.121	
LEVERAGE	-0.072	0.011	0.000	***	-0.074	0.011	0.000	***	-0.074	0.011	0.000	***
CAPX	-0.007	0.013	0.623		-0.007	0.013	0.623		-0.007	0.013	0.620	
T_Q	-0.002	0.000	0.000	***	-0.002	0.000	0.000	***	-0.002	0.000	0.000	***
CREDIT_RATING	-0.019	0.007	0.005	***	-0.018	0.007	0.006	***	-0.018	0.007	0.006	***
CASH HOLDING	0.006	0.005	0.196		0.006	0.005	0.181		0.006	0.005	0.182	
RETAINED EARNINGS	0.000	0.001	0.703		0.000	0.001	0.713		0.000	0.001	0.701	
R&D	0.000	0.000	0.268		0.000	0.000	0.270		0.000	0.000	0.269	
CFO	0.011	0.004	0.005	***	0.011	0.004	0.005	***	0.011	0.004	0.004	***
DIVERSITY_D	0.013	0.005	0.011	**								
HERF_ASSET					-0.041	0.011	0.000	***				
HERF_SALES									-0.038	0.012	0.001	***
R^2		0.396				0.396				0.396		

This table presents the results of dividend payout ratio and firm structure. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 60049. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5.5.2.2 Agency problem hypothesis

To test our agency problem hypothesis, we use two proxies of agency problems which are free cash flow and Tobin's Q. We divide our sample into two groups based on whether free cash flow (Tobin's Q) is higher than the industry median every year. We use two-digits Standard Industry Code (SIC) as our industry classification.

We show the results using free cash flow as the agency problem criterion in Table 5.4 Panel A. We find that among the group with high free cash flows, all three measures of diversification variable are highly significant. In other words, for firms with high agency problems, diversified firms pay higher dividends than focused firms. In addition, firms with higher diversification levels make higher dividend payouts. In Panel B, there is no significant difference in payout ratio between diversified firms and focused firms.

Using Tobin's Q as an alternative proxy for agency problem, we find similar results in Panel C and D. Overall, the results are consistent to the agency problem hypothesis that managers in diversified firms pay higher level of dividend to signal the good treatment of their investors.

Table 5.4

Dividend payout ratio and firm structure: agency hypothesis

Panel A: Group by Free-Cash Flow

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.054	0.012	0.000	***	0.055	0.012	0.000	***	0.055	0.012	0.000	***
SIZE	0.011	0.003	0.000	***	0.011	0.003	0.000	***	0.011	0.003	0.000	***
SALESGROWTH	-0.007	0.004	0.074	*	-0.007	0.004	0.073	*	-0.007	0.004	0.072	*
LEVERAGE	-0.043	0.014	0.003	***	-0.044	0.014	0.002	***	-0.044	0.014	0.002	***
CAPX	-0.054	0.018	0.003		-0.054	0.018	0.003	***	-0.054	0.018	0.003	***
T_Q	-0.005	0.001	0.000	***	-0.005	0.001	0.000	***	-0.005	0.001	0.000	***
CREDIT_RATING	0.016	0.006	0.010	***	0.016	0.006	0.009	***	0.016	0.006	0.010	***
CASH HOLDING	0.008	0.008	0.333		0.008	0.008	0.334		0.007	0.008	0.339	
RETAINED EARNINGS	0.002	0.001	0.114		0.002	0.001	0.106		0.002	0.001	0.106	
R&D	0.000	0.000	0.683		0.000	0.000	0.666		0.000	0.000	0.676	
CFO	0.019	0.007	0.007		0.019	0.007	0.007	***	0.019	0.007	0.007	***
DIVERSITY_D	0.014	0.006	0.028	**								
HERF_ASSET					-0.037	0.013	0.006	***				
HERF_SALES									-0.036	0.014	0.009	***
R^2		0.471				0.471				0.471		

This table presents the results by agency problem. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 34003. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.4

Dividend payout ratio and firm structure: agency hypothesis

Panel B: Group by Free-Cash Flow

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.030	0.007	0.000	***	0.030	0.007	0.000	***	0.030	0.007	0.000	***
SIZE	0.021	0.004	0.000	***	0.021	0.004	0.000	***	0.021	0.004	0.000	***
SALESGROWTH	-0.003	0.002	0.145		-0.003	0.002	0.129		-0.003	0.002	0.138	
LEVERAGE	-0.063	0.019	0.001	***	-0.066	0.019	0.001	***	-0.065	0.019	0.001	***
CAPX	0.036	0.027	0.172		0.036	0.027	0.171		0.037	0.027	0.169	
T_Q	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***
CREDIT_RATING	0.011	0.011	0.293		0.011	0.011	0.297		0.011	0.011	0.299	
CASH HOLDING	0.004	0.006	0.561		0.004	0.006	0.512		0.004	0.006	0.519	
RETAINED EARNINGS	-0.002	0.001	0.022	**	-0.002	0.001	0.021	**	-0.002	0.001	0.022	**
R&D	0.000	0.000	0.556		0.000	0.000	0.554		0.000	0.000	0.555	
CFO	0.008	0.006	0.157		0.008	0.006	0.143		0.008	0.006	0.145	
DIVERSITY_D	0.006	0.009	0.521									
HERF_ASSET					-0.039	0.022	0.076	*				
HERF_SALES									-0.033	0.023	0.143	
R^2		0.390				0.390				0.390		

This table presents the results by agency problem. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 21158. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.4

Dividend payout ratio and firm structure: agency hypothesis

Panel C: Group by Tobin's Q

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.100	0.011	0.000	***	0.100	0.011	0.000	***	0.101	0.011	0.000	***
SIZE	0.020	0.003	0.000	***	0.019	0.003	0.000	***	0.019	0.003	0.000	***
SALESGROWTH	0.001	0.003	0.662		0.001	0.003	0.693		0.001	0.003	0.682	
LEVERAGE	-0.106	0.015	0.000	***	-0.108	0.015	0.000	***	-0.108	0.015	0.000	***
CAPX	0.009	0.023	0.701		0.009	0.023	0.692		0.009	0.023	0.693	
T_Q	-0.030	0.004	0.000	***	-0.029	0.004	0.000	***	-0.029	0.004	0.000	***
CREDIT_RATING	0.025	0.006	0.000	***	0.026	0.006	0.000	***	0.025	0.006	0.000	***
CASH HOLDING	-0.008	0.012	0.510		-0.007	0.012	0.536		-0.007	0.012	0.540	
RETAINED EARNINGS	-0.004	0.001	0.000		-0.004	0.001	0.000	***	-0.004	0.001	0.000	***
R&D	0.000	0.000	0.943		0.000	0.000	0.941		0.000	0.000	0.941	
CFO	0.015	0.008	0.070		0.015	0.008	0.067	*	0.015	0.008	0.069	*
DIVERSITY_D	0.011	0.006	0.055	*								
HERF_ASSET					-0.039	0.014	0.004	***				
HERF_SALES									-0.035	0.014	0.010	**
R^2		0.37				0.37				0.37		

This table presents the results by agency problem. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 37407. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.4
Dividend payout ratio and firm structure: agency hypothesis

Panel D: Group by Tobin's Q

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	-0.001	0.006	0.861		-0.001	0.006	0.861		-0.001	0.006	0.860	
SIZE	-0.002	0.004	0.685		-0.002	0.004	0.656		-0.002	0.004	0.649	
SALESGROWTH	-0.004	0.002	0.065	*	-0.004	0.002	0.063	*	-0.004	0.002	0.062	*
LEVERAGE	0.040	0.024	0.098	*	0.039	0.024	0.104		0.039	0.024	0.107	
CAPX	-0.047	0.023	0.039	**	-0.047	0.023	0.039	**	-0.047	0.023	0.039	**
T_Q	0.001	0.001	0.248		0.001	0.001	0.242		0.001	0.001	0.241	
CREDIT_RATING	-0.011	0.009	0.225		-0.011	0.009	0.224		-0.011	0.009	0.224	
CASH HOLDING	0.004	0.005	0.445		0.004	0.005	0.435		0.004	0.005	0.430	
RETAINED EARNINGS	0.003	0.001	0.003	***	0.003	0.001	0.003	***	0.003	0.001	0.004	***
R&D	0.000	0.000	0.466		0.000	0.000	0.465		0.000	0.000	0.464	
CFO	0.006	0.005	0.245		0.006	0.005	0.242		0.006	0.005	0.237	
DIVERSITY_D	0.004	0.011	0.707									
HERF_ASSET					-0.015	0.024	0.525					
HERF_SALES									-0.018	0.024	0.442	
R^2		0.53				0.53				0.53		

This table presents the results by agency problem. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 18181. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5.5.2.3 Financial Constraint hypothesis

We include two measures of financial constraints, coverage ratio and cash flow volatility, in Table 5.5 to proceed our internal capital market hypothesis.

We use coverage ratio and cash flow volatility to measure the firm's financial constraints. We use coverage ratio to measure if firms are internally constraint as it directly measures the availability of internal funds which is consistent with several prior literatures (Cleary 1999, and Guariglia 2008). We define coverage ratio as net income dividend by interest expenses. A high coverage ratio indicates sufficient internal funds. We select firms as internally constraint firms if their coverage ratio is below the median coverage ratio with the same industry membership (2-digits SIC codes) in yearly basis.

Cash flow volatility is another common measure of financial constraint since higher cash flow volatility implies that firms are more likely to experience cash shortfall internally. Minton and Schrand (1999) shows that cash flow volatility is associated with lower investment and higher cost of accessing external financing. We construct the cash flow volatility as the standard deviation of firm's past 5 years' cash flows from operating activities. We select the firms into high financial constraint group in yearly basis as if the firms cash flow volatility is below the median cash flow volatility by the same industry membership (2-digit SIC code) and firms belongs to the low financial constraint groups if the cash flow volatility is above this industry median benchmark.

In Panel A and B, we present the results of a firm's structure's relation to dividend payout in the financially constrained sub-group and the unconstrained sub-group. We see the diversity dummy variable and two alternative diversity level variables are highly significant for the firms with financial constraint, but not in the unconstrained firms.

Similar findings are found in Panel C and D using cash flow volatility as a criterion of financial constraint. Overall, the results suggest that among the firms with the financial constraint condition, diversified firms can pay higher dividends to earnings than focused firms. This is consistent with our hypothesis that diversified firms can pay higher dividends through their efficient internal capital market when they are financially constrained.

Table 5.5

Dividend payout ratio and firm structure: financial constraint hypothesis

Panel A: Group by coverage ratio

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.001	0.010	0.909		0.001	0.010	0.903		0.001	0.010	0.906	
SIZE	0.005	0.004	0.155		0.005	0.004	0.171		0.005	0.004	0.152	
SALESGROWTH	-0.003	0.003	0.239		-0.003	0.003	0.238		-0.003	0.003	0.241	
LEVERAGE	-0.078	0.019	0.000	***	-0.079	0.019	0.000	***	-0.078	0.019	0.000	***
CAPX	-0.052	0.025	0.038		-0.052	0.025	0.037	**	-0.052	0.025	0.038	**
T_Q	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***
CREDIT_RATING	0.013	0.008	0.115		0.013	0.008	0.114		0.013	0.008	0.114	
CASH HOLDING	0.005	0.007	0.476		0.005	0.007	0.473		0.005	0.007	0.486	
RETAINED EARNINGS	0.000	0.001	0.835		0.000	0.001	0.834		0.000	0.001	0.846	
R&D	0.000	0.000	0.261		0.000	0.000	0.261		0.000	0.000	0.261	
CFO	0.023	0.007	0.001		0.023	0.007	0.001	***	0.023	0.007	0.001	***
DIVERSITY_D	0.007	0.008	0.428									
HERF_ASSET					-0.021	0.019	0.268					
HERF_SALES									-0.012	0.018	0.521	
R^2		0.384				0.384				0.384		

This table presents the results by financial constraint problem. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 38144. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.5

Dividend payout ratio and firm structure: financial constraint hypothesis

Panel B: Group by coverage ratio

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.031	0.006	0.000	***	0.031	0.006	0.000	***	0.031	0.006	0.000	***
SIZE	0.019	0.003	0.000	***	0.019	0.003	0.000	***	0.019	0.003	0.000	***
SALESGROWTH	-0.004	0.002	0.036	**	-0.004	0.002	0.035	**	-0.004	0.002	0.037	**
LEVERAGE	-0.036	0.014	0.013	**	-0.037	0.014	0.009	***	-0.037	0.014	0.010	***
CAPX	-0.013	0.016	0.431		-0.013	0.016	0.441		-0.013	0.016	0.438	
T_Q	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***
CREDIT_RATING	0.013	0.007	0.068	*	0.013	0.007	0.067	*	0.013	0.007	0.068	*
CASH HOLDING	0.008	0.006	0.187		0.008	0.006	0.180		0.008	0.006	0.176	
RETAINED EARNINGS	-0.001	0.001	0.233		-0.001	0.001	0.240		-0.001	0.001	0.239	
R&D	0.000	0.000	0.130		0.000	0.000	0.128		0.000	0.000	0.126	
CFO	0.006	0.005	0.269		0.006	0.005	0.273		0.006	0.005	0.262	
DIVERSITY_D	0.016	0.006	0.012	**								
HERF_ASSET					-0.043	0.014	0.003	***				
HERF_SALES									-0.042	0.015	0.004	***
R^2		0.46				0.46				0.46		

This table presents the results by financial constraint problem. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 17025. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.5

Dividend payout ratio and firm structure: financial constraint hypothesis

Panel C. Group by cash flow volatility

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.031	0.012	0.011	**	0.031	0.012	0.011	**	0.031	0.012	0.011	**
SIZE	0.020	0.006	0.001	***	0.020	0.006	0.001	***	0.020	0.006	0.001	***
SALESGROWTH	-0.001	0.004	0.885		0.000	0.004	0.903		0.000	0.004	0.906	
LEVERAGE	-0.041	0.026	0.116		-0.041	0.026	0.114		-0.041	0.026	0.119	
CAPX	-0.011	0.036	0.768		-0.011	0.036	0.768		-0.011	0.036	0.769	
T_Q	-0.002	0.001	0.018	**	-0.002	0.001	0.018	**	-0.002	0.001	0.018	**
CREDIT_RATING	0.010	0.013	0.416		0.010	0.013	0.417		0.010	0.013	0.417	
CASH HOLDING	-0.013	0.010	0.206		-0.013	0.010	0.204		-0.013	0.010	0.201	
RETAINED EARNINGS	-0.002	0.001	0.206		-0.002	0.001	0.209		-0.002	0.001	0.210	
R&D	0.000	0.000	0.370		0.000	0.000	0.370		0.000	0.000	0.373	
CFO	0.013	0.010	0.187		0.013	0.010	0.188		0.013	0.010	0.188	
DIVERSITY_D	0.010	0.013	0.437									
HERF_ASSET					-0.020	0.028	0.479					
HERF_SALES									-0.016	0.030	0.588	
R^2		0.405				0.405				0.405		

This table presents the results by financial constraint problem. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 41294. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.5

Dividend payout ratio and firm structure: financial constraint hypothesis

Panel D. Group by cash flow volatility

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.043	0.012	0.000	***	0.044	0.012	0.000	***	0.044	0.012	0.000	***
SIZE	0.021	0.004	0.000	***	0.020	0.004	0.000	***	0.020	0.004	0.000	***
SALESGROWTH	-0.005	0.004	0.228		-0.005	0.004	0.220		-0.005	0.004	0.228	
LEVERAGE	-0.062	0.018	0.000	***	-0.065	0.018	0.000	***	-0.065	0.018	0.000	***
CAPX	-0.069	0.025	0.006	***	-0.068	0.025	0.007	***	-0.069	0.025	0.007	***
T_Q	-0.008	0.001	0.000	***	-0.008	0.001	0.000	***	-0.008	0.001	0.000	***
CREDIT_RATING	0.009	0.008	0.292		0.009	0.008	0.281		0.008	0.008	0.295	
CASH HOLDING	0.011	0.011	0.330		0.011	0.011	0.312		0.011	0.011	0.323	
RETAINED EARNINGS	-0.001	0.001	0.338		-0.001	0.001	0.353		-0.001	0.001	0.350	
R&D	0.000	0.000	0.179		0.000	0.000	0.182		0.000	0.000	0.183	
CFO	0.018	0.009	0.054	*	0.018	0.009	0.054	*	0.018	0.009	0.055	*
DIVERSITY_D	0.018	0.007	0.008	***								
HERF_ASSET					-0.056	0.016	0.000	***				
HERF_SALES									-0.051	0.016	0.001	***
R^2		0.341				0.341				0.341		

This table presents the results by financial constraint problem. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 14538. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5.6. Robustness tests and additional analyses

Firm diversification can be an endogenous choice by a firm (Campa and Kedia 2002, Graham et al., 2002). In this section, we introduce three econometric methods to control for any endogeneity problem in our regression tests. These are the propensity score matching method and the Heckman (1979) two-stage estimation and fixed effect model.

5.6.1. Methodology

In this section, we present the estimation of the probability of diversifying. The results are reported in Table 6 and will be used in the propensity score matching method and Heckman's self-selection model.

5.6.2. Propensity-score matching

We use propensity score matching model to control for sample selection bias due to observable characteristics. As noted earlier in this thesis, Roberts and Whited (2012) comment that propensity score matching is the most commonly used methodology to address endogeneity concerns due to its simplicity. Based on the one-equation system, the key benefit of using the matched sample for the regression analysis is that it avoids specification of the function Subrahmanyam, Tang, and Wang (2014). While the disadvantage is the that the potential outcomes are taken as independent of the treatment assignment is untestable. Therefore, we use propensity score matching as a robust test to control for the difference of dividend payout accompanying firm diversification. The propensity score match method conducts two stages to estimate the average treatment effect for the treated group (ATT). We define the treatment indicator equals 1 firm diversified firms and 0 for focused firms. The outcome variable is the dividend payout ratio. In the first stage, we predict the probability of diversifying by using a logit regression, where the dependent variable takes value of 1 for diversified firms and 0 for focused firms. We use all the control variables that we use in our main regression as the independent variables in the logit regression. In the second step, based on the propensity scores obtained from the predicted probabilities in stage one, we construct the common support sample (matched sample) by using the nearest neighbourhood matching. This matching is a one-to-one matching with replacement so that one diversified firm is matched with a focused firm with the nearest propensity score. Table 5.6 Panel A presents the statistic difference between diversified firms and focused firms using the matched sample. Compared with the unmatched sample, the mean difference of control variables between diversified firms and focused firms becomes much smaller and statistically insignificant. This indicates that our propensity score match tests are generally successful. In addition, we find that the diversified firms have significantly higher dividend payout than focused firm under this propensity matched sample.

Table 5.6

The effect of diversification on dividend payout: propensity score matching estimations

Panel A: matched sample description

	Diversified firm	Focused firm (Control firms)	Difference	
	Mean	Mean	Mean	
DIV PAYOUT	0.228	0.167	0.061	***
ROA	0.027	0.021	0.006	*
SIZE	5.419	5.007	0.412	
SALESGROWTH	0.13	0.127	0.003	
LEVERAGE	0.454	0.456	-0.002	
CAPX	0.076	0.078	-0.002	**
T_Q	0.896	0.937	-0.041	*
CRD	0.526	0.492	0.034	
CASH HOLDING	0.071	0.076	-0.005	
RETAINED EARNINGS	0.082	0.039	0.043	*
CFO	-0.039	-0.048	0.009	**

This table reports the statistic difference between diversified firms and focused firms using the propensity matched sample to estimate the average treatment effect for the treated group. The treatment indicator equals 1 for diversified firms and 0 for focused firms. Detailed propensity score matched steps are described in the main content. Difference is the mean difference between diversified firms and the focused firms (the control group) using the paired T-tests. Variable definitions are provided in the notes to Table 5.1. Diversified firm is defined as the firms with multiple business segments while focused firms operate only one business sectors. There are 19721 matches. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.6 Panel B re-examines our tests of firm diversification and dividend payout ratio using the propensity-matched sample. For all three specifications of diversification measurements, diversified firms have significantly higher dividend payout ratio than focused firms. In Appendix 5.1, we find that in both criteria of agency problems using free cash flow (results reported in Panel A and B) and Tobin's Q (results reported in Panel C and D), diversified firms pay higher dividend relative to their earnings comparing to focused firms when these firms are in the agency problem groups and there is no difference in firm structure and dividend payout ratios when firms are not suffering agency problems.

In the rest of the tables of Appendix 5.1, we test our efficient internal capital market hypothesis by classifying whether firms are suffering financial constraints. Firms with higher CFO volatility (results reported in Panel E and F) or coverage ratio (results reported in Panel G and H) have higher possibility of suffering financial constraints. Results suggest that under the sample of financial constraints firms, diversified firms pay higher dividends relative to their earnings and this difference is much weaker or insignificant under the group of financially unconstrained groups.

Overall, under the propensity matching method, using the matched sample, our findings are consistent with our main regression analysis and hypothesis.

Table 5.6

Dividend payout ratio and firm structure: double-fixed effect regression model under propensity matched sample

Panel B: Full regression

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.091	0.025	0.000	***	0.091	0.025	0.000	***	0.092	0.025	0.000	***
SIZE	0.003	0.004	0.461		0.002	0.004	0.560		0.002	0.004	0.578	
SALESGROWTH	0.001	0.006	0.818		0.001	0.006	0.848		0.001	0.006	0.843	
LEVERAGE	-0.127	0.022	0.000	***	-0.130	0.022	0.000	***	-0.130	0.022	0.000	***
CAPX	0.022	0.029	0.436		0.022	0.029	0.439		0.022	0.029	0.449	
T_Q	-0.010	0.002	0.000	***	-0.009	0.002	0.000	***	-0.009	0.002	0.000	***
CREDIT_RATING	-0.026	0.009	0.003	***	-0.025	0.009	0.005	***	-0.025	0.009	0.005	***
CASH HOLDING	-0.047	0.017	0.007	***	-0.046	0.017	0.008	***	-0.046	0.017	0.008	***
RETAINED EARNINGS	0.003	0.002	0.279		0.003	0.002	0.273		0.003	0.002	0.282	
R&D	0.011	0.005	0.025	**	0.012	0.005	0.019	**	0.011	0.005	0.040	**
CFO	0.021	0.014	0.150		0.020	0.014	0.153		0.021	0.014	0.139	
DIVERSITY_D	0.018	0.007	0.013	**								
HERF_ASSET					-0.054	0.015	0.000	***				
HERF_SALES									-0.055	0.016	0.001	***
R^2		0.384				0.384				0.384		

This table presents the results of dividend payout ratio and firm structure under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 39442. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5.6.3. Heckman two-stage estimation

We use Heckman (1979) two-stage estimation tests to solve the self-selection problem. In the first stage, we use a probit regression to estimate the probability of a firm choosing to diversify. The "treatment effect" is based on the difference between diversified firms and the matched focused firms with similar propensity scores of the "decision to diversify". We follow the selections of Campa and Kedia (2002) for variables used in their first-stage Heckman test (a probit model to investigate the firms' decisions to diversify). First, firms' characteristics are used in the probit model. They include current as well as past two year's firm size, EBIT, capital expenditure as the key factors for firms' decision to diversify. They also add two dummy variables whether firms are listed on major stock exchange (listed on Nasdaq, NYSE or AMEX) and whether firms belong to S&P index. Second, they include industry level variables: Fraction of Diversified Firms in the Industry is the fraction of all the firms in the industry that are diversified firms. Fraction of Industry Sales by Diversified Firms is the fraction of industry sales accounted for by diversified firms. Third, year and economic factors are also controlled. GDP is Gross Domestic Product. GDP Growth is the growth rate of Gross Domestic Product. We report the probit estimation in Table 5.7 Panel A. In the second-stage of Heckman test, based on the estimated probability of firms' choice to diversity from Panel A, we use the estimated coefficients to construct Inverse Mills Ratios as an additional control variable which include the treatment effects for firms' self-selection into diversification in our main regressions.

Table 5.7

Dividend payout ratio and firm structure: Heckman test

Panel A. First stage of Heckman estimation: probit regression

	Coef.	Std. Err.	p-value	
(Intercept)	-1.323	0.040	0.000	***
EBIT	0.261	0.032	0.000	***
EBIT.L1	0.000	0.000	0.029	**
EBIT.L2	0.056	0.018	0.002	***
CAPX.L1	-0.001	0.003	0.689	
CAPX.L2	-0.033	0.014	0.020	**
CAPX	-0.846	0.063	0.000	***
MSE	0.027	0.012	0.033	**
GDPG	0.211	0.378	0.578	
GDP.L1	-0.006	0.000	0.000	***
SP	0.088	0.017	0.000	***
PNDIV	0.260	0.027	0.000	***
PSDIV	2.570	0.048	0.000	***
SIZE.L2	0.018	0.008	0.030	**
SIZE	0.043	0.009	0.000	***
SIZE.L1	0.068	0.011	0.000	***

This table presents probit estimates of the first stage of Heckman test. The dependent variable is firm diversification which is 1 if firms have more than one segments and 0 otherwise. EBIT is the earnings before interests and taxes deflated by total assets. CAPX is capital expenditure to assets. MSE is dummy variable which equal to 1 if the firm is listed on Nasdaq, NYSE, or AMEX, and 0 otherwise. GDP and GDP are Gross Domestic Product and the growth rate of GDP. SP is a dummy variable taking value of 1 if the firms is part of S&P index and 0 otherwise. PNDIV is the fraction of diversified firms in the industry is the fraction of all firms in the industry that are diversified firms. PSDIV is the fraction of industry sales accounted for by diversified firms. The number of observation is 64331. Standard errors are clustered by firms. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

We repeat our full tests of relationship between firm diversification and dividend payout ratio by including this Inverse Mills Ratio as a control for selection bias. In Table 5.7 Panel B, Consistent with our earlier findings, we find that diversified firms have significantly higher dividend payout ratio compared to focused firms.

In Appendix 5.2, we also test our analysis by dividing firms into groups based on whether they have agency problems. For Panel A and B, we use free cash flow as the criteria of agency problem. For Panel C and D, we use Tobin's Q as the agency problem proxy. Under the second stage of Heckman test, we find that diversified firms pay in a significantly higher dividend payout ratio than focused firms under the agency problem groups. In addition, there is no difference in dividend payout ratio in the groups with no agency problems.

Table 5.7

Dividend payout ratio and firm structure: double-fixed effect regression model with second stage of Heckman test

Panel B. Full regression

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.031	0.009	0.001	***	0.031	0.009	0.001	***	0.031	0.009	0.001	***
SIZE	0.019	0.003	0.000	***	0.018	0.003	0.000	***	0.018	0.003	0.000	***
SALESGROWTH	-0.003	0.003	0.309		-0.003	0.003	0.306		-0.003	0.003	0.313	
LEVERAGE	-0.062	0.016	0.000	***	-0.064	0.016	0.000	***	-0.064	0.016	0.000	***
CAPX	-0.032	0.019	0.082	*	-0.032	0.018	0.085	*	-0.032	0.018	0.085	*
T_Q	-0.005	0.001	0.000	***	-0.005	0.001	0.000	***	-0.005	0.001	0.000	***
CREDIT_RATING	0.011	0.008	0.151		0.011	0.008	0.143		0.011	0.008	0.149	
CASH HOLDING	-0.001	0.008	0.880		-0.001	0.008	0.904		-0.001	0.008	0.896	
RETAINED EARNINGS	-0.002	0.001	0.030	**	-0.002	0.001	0.032	**	-0.002	0.001	0.032	**
R&D	0.000	0.000	0.176		0.000	0.000	0.179		0.000	0.000	0.180	
CFO	0.020	0.006	0.002	***	0.020	0.006	0.002	***	0.020	0.006	0.002	***
INVERSE MILLS RATIO	-0.004	0.006	0.512		-0.004	0.006	0.531		-0.004	0.006	0.508	
DIVERSITY_D	0.019	0.006	0.004	***								
HERF_ASSET					-0.050	0.014	0.000	***				
HERF_SALES									-0.049	0.015	0.001	***
R^2		0.361				0.361				0.361		

This table presents the results of dividend payout ratio and firm structure under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 55368. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Similar tests are summarized in Appendix 5.2 by dividing the firms into whether or not they are financially constrained. Panels E and F uses the coverage ratio and Panels G and H uses operating cash flow volatility as the criteria of financial constraints separately. Both criteria grouping methods show consistent findings. For the financially constrained firms, diversified firms pay higher dividends relative to their earnings comparing to focused firms and there is no significant difference in dividend payout ratio under the financially unconstraint groups.

For all the above analysis, the Inverse Mills Ratios are all insignificant in all the analysis and our results are robust after controlling for self-selection problems.

5.7. Additional tests

5.7.1. Alternative measures of dividend payment

We define dividend payout ratio as the total dividend scaled by net income as our dependent variable. We use different dividend payout ratios measures as in our model for robustness checks. First, in unreported tests, we define our dependent variable as total cash dividends paid to common and preferred shareholders scaled by earnings after taxes and interest but before extraordinary items. Second, we use whether to pay dividend as an additional proxy for dividend policy and alternatively, we use a logit regression model with fixed effect to test our main analysis.

Overall, we find results consistent with our main regression analyses testing our agency problem hypothesis and efficient internal capital market hypothesis.

5.7.2. Financial Life-cycle dividend theory

Prior empirical studies suggest that mature and established firms usually pay dividends (DeAngelo, DeAngelo, and Stulz 2006, Chay and Suh 2009). Firms at their mature stage usually have large cumulative profits and are, therefore, more likely to pay dividends. In addition, firms at the maturity stage tend to have lower investment opportunities and higher levels of free cash flows. This can lead to higher agency problems. We use firm age and size as two proxies for the firm's maturity. We predict that diversified firms pay higher dividends than focused firms when they are at the maturity stage.

In Table 5.8, we present the results of our main regression across the group of mature (large) firms and young (small) firms. In Panel A and B, we find that for young firms, there is no difference in the dividend payout[13] between diversified firms and among the firms that have reached maturity, diversified firms have higher dividend payout ratio than focused firms. Similarly, in Panel C and D, we find significantly higher dividend payout ratio for diversified firms than for focused firms in the large firms' sample, but no firm structure difference of the payout in the small firms' group. This is consistent with the business-life cycle theory of dividend payment, where firms pay dividends when they are at maturity since there are fewer investment opportunities than at the start-up stage. Due to the higher level of free cash flows held by the mature firms, diversified firms pay higher dividends than focused firms, in line with our main analyses.

[13] Our results are also robust using dividend payment in the logistic model as an alternative of dividend payout ratio.

Table 5.8

Dividend payout ratio and firm structure: business-cycle theory

Panel A: Group by firm age

	Old firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.028	0.014	0.039	**	0.029	0.014	0.035	**	0.029	0.014	0.035	**
SIZE	0.027	0.005	0.000	***	0.026	0.005	0.000	***	0.026	0.005	0.000	***
SALESGROWTH	0.001	0.006	0.849		0.001	0.006	0.860		0.001	0.006	0.848	
LEVERAGE	-0.064	0.026	0.013	**	-0.067	0.026	0.009	***	-0.067	0.026	0.009	***
CAPX	-0.067	0.032	0.035	**	-0.067	0.032	0.035	**	-0.067	0.032	0.034	**
T_Q	-0.007	0.001	0.000	***	-0.007	0.001	0.000	***	-0.007	0.001	0.000	***
CREDIT_RATING	0.010	0.010	0.308		0.011	0.010	0.299		0.010	0.010	0.308	
CASH HOLDING	-0.002	0.011	0.841		-0.002	0.011	0.866		-0.002	0.011	0.846	
RETAINED EARNINGS	-0.003	0.002	0.031	**	-0.003	0.002	0.033	**	-0.003	0.002	0.035	**
R&D	0.000	0.000	0.268		0.000	0.000	0.269		0.000	0.000	0.273	
CFO	0.024	0.011	0.032	**	0.024	0.011	0.033	**	0.024	0.011	0.032	**
DIVERSITY_D	0.026	0.008	0.002	***								
HERF_ASSET					-0.065	0.017	0.000	***				
HERF_SALES									-0.063	0.017	0.000	***
R^2		0.40				0.40				0.40		

This table presents the results by firm's maturity. The regressions are presented across the groups of maturities. The criteria for agency problem are firm age (Panel A &B) and firm size (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 49093. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.8

Dividend payout ratio and firm structure: business-cycle theory

Panel B: Group by firm age

	Young firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.028	0.006	0.000	***	0.028	0.006	0.000	***	0.028	0.006	0.000	***
SIZE	0.004	0.003	0.144		0.004	0.003	0.151		0.004	0.003	0.128	
SALESGROWTH	-0.002	0.001	0.127		-0.002	0.001	0.124		-0.002	0.001	0.135	
LEVERAGE	-0.060	0.013	0.000	***	-0.060	0.013	0.000	***	-0.058	0.013	0.000	***
CAPX	0.006	0.014	0.644		0.006	0.014	0.644		0.006	0.014	0.646	
T_Q	-0.001	0.000	0.017	**	-0.001	0.000	0.016	**	-0.001	0.000	0.016	**
CREDIT_RATING	0.006	0.007	0.401		0.006	0.007	0.398		0.006	0.007	0.394	
CASH HOLDING	0.006	0.005	0.168		0.006	0.005	0.166		0.006	0.005	0.181	
RETAINED EARNINGS	0.000	0.001	0.606		0.000	0.001	0.612		0.000	0.001	0.587	
R&D	0.000	0.000	0.267		0.000	0.000	0.267		0.000	0.000	0.267	
CFO	0.006	0.004	0.105		0.006	0.004	0.103		0.006	0.004	0.111	
DIVERSITY_D	-0.003	0.006	0.618									
HERF_ASSET					0.006	0.015	0.685					
HERF_SALES									0.019	0.015	0.215	
R^2		0.215				0.215				0.215		

This table presents the results by firm's maturity. The regressions are presented across the groups of maturities. The criteria for agency problem are firm age (Panel A &B) and firm size (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 2561. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.8

Dividend payout ratio and firm structure: business-cycle theory

Panel C: Group by firm size

	Large firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.012	0.016	0.453		0.012	0.016	0.437		0.013	0.016	0.429	
SIZE	0.021	0.005	0.000	***	0.020	0.005	0.000	***	0.020	0.005	0.000	***
SALESGROWTH	-0.007	0.004	0.112		-0.007	0.004	0.113		-0.007	0.004	0.110	
LEVERAGE	-0.001	0.021	0.949		-0.004	0.021	0.865		-0.004	0.021	0.841	
CAPX	-0.092	0.029	0.001	***	-0.092	0.029	0.001	***	-0.092	0.029	0.001	***
T_Q	-0.003	0.001	0.001	***	-0.003	0.001	0.001	***	-0.003	0.001	0.001	***
CREDIT_RATING	0.023	0.010	0.018	**	0.023	0.010	0.018	**	0.023	0.010	0.019	**
CASH HOLDING	0.019	0.007	0.009	***	0.019	0.007	0.010	**	0.019	0.007	0.009	***
RETAINED EARNINGS	0.000	0.002	0.936		0.000	0.002	0.938		0.000	0.002	0.954	
R&D	0.000	0.000	0.024	**	0.000	0.000	0.025	**	0.000	0.000	0.026	**
CFO	0.026	0.008	0.001	***	0.026	0.008	0.001	***	0.026	0.008	0.001	***
DIVERSITY_D	0.023	0.008	0.004	***								
HERF_ASSET					-0.056	0.018	0.001	***				
HERF_SALES									-0.055	0.018	0.002	***
R^2		0.381				0.381				0.381		

This table presents the results by firm's maturity. The regressions are presented across the groups of maturities. The criteria for agency problem are firm age (Panel A &B) and firm size (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation 28283. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5.8

Dividend payout ratio and firm structure: business-cycle theory

Panel D: Group by firm size

	Small firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.036	0.006	0.000	***	0.036	0.006	0.000	***	0.036	0.006	0.000	***
SIZE	0.010	0.003	0.000	***	0.010	0.003	0.000	***	0.010	0.003	0.000	***
SALESGROWTH	-0.002	0.002	0.291		-0.002	0.002	0.281		-0.002	0.002	0.298	
LEVERAGE	-0.070	0.012	0.000	***	-0.071	0.012	0.000	***	-0.070	0.012	0.000	***
CAPX	0.011	0.014	0.447		0.011	0.014	0.440		0.011	0.014	0.443	
T_Q	-0.002	0.001	0.001	***	-0.002	0.001	0.001	***	-0.002	0.001	0.001	***
CREDIT_RATING	0.000	0.005	0.961		0.000	0.005	0.960		0.000	0.005	0.965	
CASH HOLDING	0.003	0.006	0.652		0.003	0.006	0.625		0.003	0.006	0.654	
RETAINED EARNINGS	0.000	0.001	0.527		0.000	0.001	0.528		0.000	0.001	0.532	
R&D	0.000	0.000	0.000	***	0.000	0.000	0.000	***	0.000	0.000	0.000	***
CFO	0.003	0.005	0.494		0.003	0.005	0.485		0.003	0.005	0.493	
DIVERSITY_D	0.008	0.006	0.188									
HERF_ASSET					-0.028	0.014	0.050	*				
HERF_SALES									-0.019	0.014	0.170	
R^2		0.384				0.384				0.384		

This table presents the results by firm's maturity. The regressions are presented across the groups of maturities. The criteria for agency problem are firm age (Panel A &B) and firm size (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 27862. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5.8. Conclusion

We document evidence that firm diversification is an important determinant of payout policy. Specifically, we analyze the effects of the firm diversification on firm's dividend policy using 10,000 firm-year observations for diversified and focused firms from 1980 to 2017. Furthermore, we introduce two hypotheses deviated from firm diversification to support our analyses.

Our main finding suggests that diversified firms pay, on average, higher dividends comparing to focused firms, which corroborates the substitute hypothesis of La Porta, Lopez-de-Silanes, Shleifer, and Vishny (2000). To explain our main findings, we proceed with two additional tests. First, we group our sample by agency. We find that among the firms with agency problems, diversified firms pay significantly higher dividends than focused firms which confirms the substitute hypothesis in our main finding. Second, we group firms according to their financial constraints. Our results suggest that among the financially constrained group, diversified firms pay higher dividends than focused firms, which lends s support to the hypothesis that an efficient internal capital market affects dividend policy.

Our work calls attention to the role played by firm structure in payout policy. It aligns with empirical studies of the diversification discount puzzle, especially those which try to solve this puzzle by removing the endogeneity of self-selection (see Campa and Kedia 2002, Graham et al., 2002). Our findings are consistent with Villalonga's (2004), of significant impact of dividend payment on firms' decisions to diversify. Although Villalonga (2004) implicitly inserts dividend payment as one of the additional variables in the self-selection model of Campa and Kedia (2002), no explanations are given. Our empirical study provides solid evidence of why dividend payments are important in a firm's diversification.

6. Conclusion

In this section, we summarize the findings and contributions of three essays.

6.1. Main findings

The principal findings of this dissertation are threefold. In the first essay, we show that median forecasting based on quantile regression performs significantly better than mean forecasting based on OLS in predicting a firm's profitability. Our forecasting comparison is based on actual profitability data including gross profitability (GP), operating profit (one common method, two methods adjusted for accruals following Ball et al., 2016), RNOA, ROE, and ROA. The finding is robust to the use of different models: the forecasting model by parsimonious AR (1) model, the model of Fairfield et al. (2009) and both the earning's decomposition models of Sloan (1996). The results still hold for different forecasting periods: pre financial crisis and post financial crisis. We confirm this finding by performing long-term forecasting from 2 years ahead forecasting to 5 years ahead forecasting. Overall, our results suggest that median forecasting based on quantile regression delivers a better central tendency than mean forecasting based on OLS regression.

Our findings shed light on those of Schröder and Yim (2017). We confirm that an industry-specific model is more accurate than an economy-wide model if we use a broader industry identification code such as the 2-digit SIC and Fama and French 12 Industry classification. Compared to a narrowly defined industry classification, broad industry classification produces more accurate in-sample estimation with sufficient number of observations. In the unreported analyses, we replicate the findings of Fairfield et al. (2009) of no incremental advantages from using an industry-specific model over economy-wide model, using their industry classification. Overall, the implication is that mean reversion can be different across industries. By using an industry-specific forecasting model with broader industry classification, we can obtain more accurate profitability forecasting than using the economy-wide forecasting model.

We perform a hedge portfolio analysis based on the information advantage of quantile regression over OLS regression. We sort stocks based on the difference of predicted profitability based on quantile regression and OLS regression models and construct the hedged portfolios. For all profitability measures used in the analyses, there are positive and significant abnormal returns from the portfolios. This means that the market participants do not fully price stocks by using the incremental information picked up by quantile regression. Our findings are robust to various stock screening criteria.

In essay 2, we investigate the relationship between firm diversification and loss reversal profitability. Loss reversal indicates a special perspective on a firm's performance which is that loss-making firms become profitable in the following year. We find empirically a significant and positive relationship between the loss reversal profitability and the firm's diversity level. Inspired by the real option theory raised by Hayn (1995), our study attempts to use an abandonment option to explain the difference of loss reversal profitability difference between firms with different structures. Our results suggest that diversified firms have significantly higher loss reversal profitabilities through exercising their abandonment options. This is consistent with the hypothesis that diversified firms hold a selection of options allowing them to abandon loss-making assets or segment more efficiently than focused firms. In additional work, we find results fitting agency theory where, among the firms with high agency problems, managers are reluctant to abandon their loss-making business lines because of a desire to maintain professional reputations. This reduces the effect of exercising abandonment options, between diversified firms and focused firms. These results are consistent after controlling for endogeneity problems by using propensity score matching, Heckman test and two-way fixed effect model. We also demonstrate robustness by using different proxies of agency problems and our our definitions ofabandonment option measures.

In essay three, we conduct cross-sectional analyses by investigating the relation between firm diversification and dividend payout policy. Overall, diversified firms have significantly higher dividend payout ratios than focused firms. To push this further, we test two hypotheses: an agency problem hypothesis and an efficient internal market hypothesis. On the one hand, we find that diversified firms pay much higher dividends than focused firms among those firms with agency problems, which is consistent with the substitute hypothesis raised by La Porta, Lopez-de-Silanes, Shleifer and Vishny (2000). On the other hand, we find that diversified firms pay higher dividends than focused firms among firms with financial constraints. Overall, the results are consistent with diversified firms paying higher dividends than focused firms because of a necessity to reduce the agency problem. Efficient internal capital markets within diversified firms can be used to distribute dividends if firms are externally constrained. Our results are robust to different agency problem proxies and financial constraint measures. Several econometric methodologies are applied to our analyses.

6.2. Contributions

Summing up, under three inter-related empirical investigations, we make the following contributions.

In Essay One we address the importance of median forecast based on quantile regression, especially forecasting a sample with skewness such as firm's profitability. First, we shed light on the importance of quantile regression as a counterpart of OLS regression in forecasting firm's profitability, constructing forecasting models, using various profitability measures with short term and long-term analyses in demonstrating the better forecasting accuracy of quantile regression. In addition, we show that positive abnormal returns can be generated through hedge portfolio analyses based on the incremental advantages of quantile regression. This confirms that market participants do not fully impound the information contained in the quantile regression.

In Essay Two, we contribute to the literature on the performance of loss-making firms by showing how firm structure (diversified vs. focused firms) is related to loss reversal probabilities. Particularly, we demonstrate how important a firm's structure is, related to the firm's strategy and performance when suffering losses. In addition, we consider how the efficiency of exercising an abandonment option might change when agency problem occurs among the managers. Finally, we propose several robustness checks on our findings, including the new abandonment option proxies we created.

In Essay Three, we explore the dividend puzzle by demonstrating how firm structure is related to dividend policy. We generalize the main tests as well as identifying the incentives that cause the difference of dividend policy by firm structure. Finally, we conduct empirical tests, controlling for endogeneity problems, and propose robustness checks on our findings.

References

Akbas, F., Jiang, C., Koch, P.D., 2017. The Trend in Firm Profitability and the Cross Section of Stock Returns. Accounting Review. doi:10.2308/accr-51708

Allen, D.E., Salim, H.M., 2005. Forecasting profitability and earnings: a study of the UK market (1982–2000). Applied Economics 37, 2009–2018. doi:10.1080/00036840500244634

Almeida, H., Campello, M., Weisbach, M, S., 2004. The Cash Flow Sensitivity of Cash. Journal of Finance 59, 1777–1804. doi:10.1111/j.1540-6261.2004.00679.x

Ambarish, R., John, K., Williams, J., 1987. Efficient Signaling with Dividends and Investments. Journal of Finance 42, 321-343. doi:10.1111/j.1540-6261.1987.tb02570.x

Amihud, Y., Lev, B., 1981. Risk Reduction as a Managerial Motive for Conglomerate Mergers. Bell Journal of Economics 12, 605-17. doi:10.2307/3003575.

Asquith, P., Mikhail, M.B., Au, A.S., 2005. Information content of equity analyst reports. Journal of Financial Economics 75, 245–282. doi:10.1016/j.jfineco.2004.01.002

Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V., 2016. Accruals, cash flows, and operating profitability in the cross section of stock returns. Journal of Financial Economics 121, 28–45. doi:10.1016/j.jfineco.2016.03.002

Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V.V., 2015. Deflating profitability. Journal of Financial Economics 117, 225–248. doi:10.1016/j.jfineco.2015.02.004

Barber, B.M., Lyon, J.D., 1996. Detecting abnormal operating performance: The empirical power and specification of test statistics. Journal of Financial Economics 41, 359–399. doi:16/0304-405X(96)84701-5

Bassett Jr, G.W., Chen, H.-L., 2001. Portfolio style: Return-based attribution using quantile regression. Empirical Economics 26, 293–305. doi:10.1007/s001810100074

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. Journal of Accounting and Economics, Properties of Accounting Earnings 24, 3–37. doi:10.1016/S0165-4101(97)00014-1

Bates, T, W., Lemmon, M, L., Bizjak, J, M., 1997. Pay-for-Performance and Firm Diversification. Available at SSRN: https://ssrn.com/abstract=41841

Beaver, W., McNichols, M., Price, R., 2007. Delisting returns and their effect on accounting-based market anomalies. Journal of Accounting and Economics 43, 341–368.

Benartzi, S., Michaely, R., Thaler, R., 1997. Do Changes in Dividends Signal the Future or the Past? Journal of Finance 52, 1007-1034. doi:10.1111/j.1540-6261.1997.tb02723.x

Berger, P, G., Ofeck, E., 1995. Diversification's effect on firm value. Journal of Financial Economics 37, 39-65. 10.1016/0304-405X(94)00798-6

Berger, P., Ofek, E., Swary, I., 1996. Investor Valuation of the Abandonment Option. Journal of Financial Economics 42, 257–287.

Bergh, D, D., Johnson, R, A., Dewitt, R, L., 2008. Restructuring Through Spin-off or Sell-off: Transforming Information Asymmetries into Financial Gain. Strategic Management Journal 29, 133–148. doi:10.1002/smj.652

Bernheim, B., Whinston, M., 1990. Multimarket Contact and Collusive Behavior. The RAND Journal of Economics, 21, 1-26.

Bhattacharya, S., Ritter, J, R., 1983. Innovation and communication: Signalling with partial disclosure, Review of Economic Studies 50, 331-346

Bhide, A., 1990. Reversing Corporate Diversification. Journal of Applied Corporate Finance 3, 70-81. doi:10.1111/j.1745-6622.1990.tb00201.x

Bodnar, G.M., Tang, C., Weintrop, J., 1999. Both Sides of Corporate Diversification: The

Value Impacts of Geographic and Industrial Diversification. Unpublished manuscript, Paul H. Nitze School of Advanced International Studies, Johns Hopkins University.

Boot, A. W. A. 1992. Why Hang on to Losers? Divestitures and Takeovers. Journal of Finance 47, 1401–1423. doi:10.1111/j.1540-6261.1992.tb04663.x

Bou, J.C., Satorra, A., 2007. The Persistence of Abnormal Returns at Industry and Firm Levels: Evidence from Spain. Strategic Management Journal 28, 707–722.

Bowen, R., Ducharme, L., Shores, D., 1995. Stakeholders' Implicit Claims and Accounting Method Choice. Journal of Accounting and Economics 20, 255–95.

Brooks, L.D., Buckmaster, D.A., 1976. Further evidence of the time series properties of accounting income. Journal of Finance 31, 1359–1373. doi:10.1111/j.1540-6261.1976.tb03218.x

Brown, P., Ball, R., 1967. Some Preliminary Findings on the Association between the Earnings of a Firm, Its Industry, and the Economy. Journal of Accounting Research 5, 55. doi:10.2307/2489908

Burgstahler, D., Dichev, I., 1997. Earnings Management to Avoid Earnings Decreases and Losses. Journal of Accounting and Economics 24, 99-126, ISSN 0165-4101, https://doi.org/10.1016/S0165-4101(97)00017-7.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Campa, J, M., Kedia, S., 2002. Explaining the Diversification Discount. Journal of Finance 57, 1731–1762. doi:10.1111/1540-6261.00476

Campbell, J.L., 2015. The Fair Value of Cash Flow Hedges, Future Profitability, and Stock Returns. Contemporary Accounting Research 32, 243–279. doi:10.1111/1911-3846.12069

Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52, 57–82. doi:10.1111/j.1540-6261.1997.tb03808.x

Chang, W.-J., Monahan, S., Ouazad, A., Vasvari, F., 2016. The Higher Moments of Future Earnings. *Working paper*

Chen, K., Ying, Z., Zhang, H., Zhao, L., 2008. Analysis of least absolute deviation. Biometrika 95, 107–122. doi:10.1093/biomet/asm082

Chen, S, S., 2006. The Economic Impact of Corporate Capital Expenditures: Focused Firms versus Diversified Firms. Journal of Financial and Quantitative Analysis 41, 341-355.doi:10.1017/S002210900000209X

Chen, Z, J., Rey, P., 2012. Loss Leading as an Exploitative Practice. American Economic Review 102, 3462-482. http://www.jstor.org/stable/41724641.

Cheng, Q., 2005. What Determines Residual Income? Accounting Review 80, 85–112.

Chevalier, J., 1999. Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms. Unpublished working paper, University of Chicago.

Comment, R., Jarrell, G, A., 1995. Corporate Focus and Stock Returns, Journal of Financial Economics 37, 67-87, ISSN 0304-405X, https://doi.org/10.1016/0304-405X(94)00777-X.

Courgeau, D., 2012. Probability and Social Science: Methodological Relationships between the two Approaches. Springer Science & Business Media.

DeAngelo, H, C., DeAngelo, L., Stulz, R., 2006. Dividend policy and the earned/contributed capital mix: A test of the life-cycle theory, Journal of Financial Economics 81, 227-254

Degeorge, F., Patel, J., Zeckhauser, R., 1999. Earnings Management to Exceed Thresholds. Journal of Business 72, 1-33. doi:10.1086/209601.

Demirakos, E.G., Strong, N.C., Walker, M., 2004. What Valuation Models Do Analysts Use? Accounting Horizons; Sarasota 18, 221–240.

Demsetz, H., 1983. The structure of ownership; Theory and consequences. Journal of Political Economy 93, 1155-1177.

Denis, D. J., Denis, D. K., Sarin, A., 1997. Agency Problems, Equity Ownership, and Corporate Diversification. Journal of Finance 52, 135–160. doi:10.1111/j.1540-6261.1997.tb03811.x

Denis, D. J., Denis, D. K., Sarin, A., 1997. Agency Problems, Equity Ownership, and Corporate Diversification. Journal of Finance 52, 135–160. doi:10.1111/j.1540-6261.1997.tb03811.x

Diebold, F., 2015. Forecasting: in Economics, Business, Finance and Beyond. Department of Economics, University of Pennsylvania.

Dimitrov, V., Tice, S., 2006. Corporate Diversification and Credit Constraints: Real Effects across the Business Cycle, Review of Financial Studies 19, 1465–1498, https://doi.org/10.1093/rfs/hhj028

Easterbrook, F., 1984. Two Agency-Cost Explanations of Dividends. American Economic Review 74, 650-659. Retrieved from http://www.jstor.org/stable/1805130

Easton, P. and Zmijewski, M. 1989. Cross-Sectional Variation in the Stock Market Response to the Announcement of Accounting Earnings. Journal of Accounting and Economics11, 117-142. http://dx.doi.org/10.1016/0165-4101(89)90003-7

Fairfield, P.M., Ramnath, S., Yohn, T.L., 2009. Do Industry-Level Analyses Improve Forecasts of Financial Performance? Journal of Accounting Research 47, 147–178.

Fama, E, F., French, K, R., 1999. The Corporate Cost of Capital and the Return on Corporate Investment. Journal of Finance 54, 1939–1967. doi:10.1111/0022-1082.00178

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56. doi:10.1016/0304-405X(93)90023-5

Fama, E.F., French, K.R., 2000. Forecasting Profitability and Earnings. Journal of Business 73, 161–75.

Fama, E.F., French, K.R., 2008. Dissecting Anomalies. Journal of Finance 63, 1653–1678. doi:10.1111/j.1540-6261.2008.01371.x

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22. doi:10.1016/j.jfineco.2014.10.010

Fama, E.F., French, K.R., 2016. Dissecting Anomalies with a Five-Factor Model. Review of Financial Studies 29, 69–103.

Fama, E.F., French, K.R., 2017. International tests of a five-factor asset pricing model. Journal of Financial Economics 123, 441–463. doi:10.1016/j.jfineco.2016.11.004

Faulhaber, G, R., 2005. Cross-subsidy Analysis with more than Two Services. Journal of Competition Law & Economics 1, 441–448, https://doi.org/10.1093/joclec/nhi014

Fernandez, P., 2002. Valuation Methods and Shareholder Value Creation. Academic Press.

Finkelstein, S., Hambrick, D., 1996. Strategic leadership: Top executives and their effects on organizations. 457. St. Paul, MN: West Publishing

Fluck, Z., 1998. Optimal Financial Contracting: Debt versus Outside Equity, Review of Financial Studies 11, 383–418, https://doi.org/10.1093/rfs/11.2.383

Fluck, Z., 1999. The Dynamics of the Management-Shareholder Conflict, Review of Financial Studies 12, 379–404, https://doi.org/10.1093/rfs/12.2.379

Freeman, R.N., Ohlson, J.A., Penman, S.H., 1982. Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation. Journal of Accounting Research 20, 639–653. doi:10.2307/2490890

Gebhardt, W.R., Lee, C.M.C., Swaminathan, B., 2001. Toward an implied cost of capital. J. Account. Res. 39, 135–176. doi:10.1111/1475-679X.00007

George, R., Kabir, R., 2008. Business Groups and Profit Redistribution: A Boon or Bane for Firms? Journal of Business Research 61, 1004-1014, ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2007.12.002.

Givoly, D., Hayn, C., 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? Journal of Accounting and Economics 29, 287–320.

Graham, J, R., Harvey, C, R., 2005. The Long-run Equity Risk Premium. Finance Research Letters 2, 185-194, ISSN 1544-6123, https://doi.org/10.1016/j.frl.2005.08.003.

Graham, J, R., Lemmon, M, L., Wolf, J, G., 2002. Does Corporate Diversification Destroy Value? Journal of Finance 57, 695–720. doi:10.1111/1540-6261.00439

Grant, R, M., Jammine, A, P., Thomas, H., 1988. Diversity, Diversification, and Profitability Among British Manufacturing Companies, 1972-1984, Academy of Management 31, 771-801

Graves, S.B., Waddock, S.A., 1990. Institutional ownership: Affect on corporate strategy. Academy of Management Executive 4, 75-83.

Gravetter, F.J., Wallnau, L.B., 2016. Statistics for The Behavioral Sciences. Cengage Learning.

Green, J., Hand, J.R.M., Zhang, X.F., 2017. The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. Review of Financial Studies doi:10.1093/rfs/hhx019

Hann, R. N., Ogneva, M. and Ozbas, O., 2013. Corporate Diversification and the Cost of Capital. Journal of Finance 68, 1961–1999. doi: 10.1111/jofi.12067

Hayn, C., 1995. The Information Content of Losses. Journal of Accounting and Economics 20, 123-153

Healey, J.F., 2015. The Essentials of Statistics: A Tool for Social Research. Cengage Learning.

Heckman, J. J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153-161

Hoechle, D., Schmid, M., Walter, I., Yermack, D., 2012. How Much of the Diversification Discount can be Explained by Poor Corporate Governance? Journal of Financial Economics, 103, 41-60, ISSN 0304-405X, https://doi.org/10.1016/j.jfineco.2011.03.025.

Hoechle, D., Schmid, M., Walter, I., Yermack, D., 2012. How Much of the Diversification Discount can be Explained by Poor Corporate Governance? Journal of Financial Economics, 103, 41-60, ISSN 0304-405X, https://doi.org/10.1016/j.jfineco.2011.03.025.

Hou, K., Xue, C., Zhang, L., 2015. Digesting Anomalies: An Investment Approach. Review of Financial Studies. 28, 650–705. doi:10.1093/rfs/hhu068

Imam, D.S., Barker, R., Clubb, C., 2008. The Use of Valuation Models by UK Investment Analysts. European Accounting Review 17, 503–535. doi:10.1080/09638180802016650

Jalali, N., Babanezhad, M., 2011. Quantile Regression due to Skewness and Outliers. Applied Mathematical Sciences 5, 1947–1951.

Jensen, M, C., 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. American Economic Review 76, 323-29. http://www.jstor.org/stable/1818789

Jensen, M, C., Murphy, K, J., 1990. Performance Pay and Top-Management Incentives. Journal of Political Economy 98, 225-64. http://www.jstor.org/stable/2937665.

Jensen, M, C., Murphy, K, J., 1990. Performance Pay and Top-Management Incentives," Journal of Political Economy 98, 225-264.

Jensen, M., Meckling, W, H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure, Journal of Financial Economics 3, 305-360.

John, K, Williams, J., 1985. Dividends, Dilution, and Taxes: A Signaling Equilibrium, Journal of Finance 40, 1053-1070. doi:10.1111/j.1540-6261.1985.tb02363.x

Joos, P., Plesko G.A., 2005. Valuing Loss Firms. Accounting Review 80, 847-870.

Journal of Political Economy 102, 1248-280. http://www.jstor.org/stable/2138786

Khanna, N., Tice, S., 2001. The Bright Side of Internal Capital Markets. Journal of Finance 56, 1489–1528. doi:10.1111/0022-1082.00377

Kim, T.-H., White, H., 2004. On more robust estimation of skewness and kurtosis. Finance Research Letters 1, 56–73. doi:10.1016/S1544-6123(03)00003-5

Koenker, R., 2015. Quantile Regression, in: Wright, J.D. (Ed.), International Encyclopedia of the Social & Behavioral Sciences (Second Edition). Elsevier, Oxford, pp. 712–718. doi:10.1016/B978-0-08-097086-8.42074-X

Koenker, R., Bassett, G., 1978. Regression Quantiles. Econometrica 46, 33–50. doi:10.2307/1913643

Konstantinidi, T., Pope, P.F., 2016. Forecasting Risk in Earnings. Contemporary Account Research 33, 487–525. doi:10.1111/1911-3846.12158

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R, W., 2000. Agency Problems and Dividend Policies around the World. Journal of Finance 55, 1–33. doi:10.1111/0022-1082.00199

Lamont, O., 1997. Cash Flow and Investment: Evidence from Internal Capital Markets. Journal of Finance 2, 83–109. doi:10.1111/j.1540-6261.1997.tb03809.x

Lamont, O., Polk, C., Requejo, J, S., 2001. Financial Constraints and Stock Returns, Review of Financial Studies 14, 529–554, https://doi.org/10.1093/rfs/14.2.529

Lang, H., P., Stulz, R, M., 1994. Tobin's Q, Corporate Diversification, and Firm Performance. Journal of Political Economy 102, 1248-280. http://www.jstor.org/stable/2138786.

Lauridsen, S., 2000. Estimation of value at risk by extreme value methods. Extremes 3, 107–144.

Lawrence, A., Sloan, R., Sun, E., 2017. Why are Losses Less Persistent than Profits? Curtailments vs. Conservatism. Management Science 63, 1-22.

Lewellen, W, G, 1971. A Pure Financial Rationale for The Conglomerate Merger, Journal of Finance 26, 521–537. doi:10.1111/j.1540-6261.1971.tb00912.x

Li, D, D., Li, S., 1996. A Theory of Corporate Scope and Financial Structure. Journal of Finance 51, 691-709. doi:10.1111/j.1540-6261.1996.tb02699.x

Li, K, K., 2011. How Well Do Investors Understand Loss Persistence? Review of Accounting Studies 16, 630-667. https://doi.org/10.1007/s11142-011-9157-4.

Li, K.K., Mohanram, P., 2014. Evaluating cross-sectional forecasting models for implied cost of capital. Rev Account Stud 19, 1152–1185. doi:10.1007/s11142-014-9282-y

Li, N., Richardson, S., Tuna, İ., 2014. Macro to micro: Country exposures, firm fundamentals and stock returns. Journal of Accounting and Economics 58, 1–20. doi:10.1016/j.jacceco.2014.04.005

Lim, C, Y., Thong, T, Y., Ding, D, K., 2008. Firm Diversification and Earnings Management: Evidence from Seasoned Equity Offerings. Review of Quantitative Finance and Accounting 30: 69-92. https://doi.org/10.1007/s11156-007-0043-x

Lipe, R., Kormendi, R., 1994. Mean reversion in annual earnings and its implications for security valuation. Review of Quantitative Finance and Accounting 4, 27–46. doi:10.1007/BF01082663

Liu, J., Nissim, D., Thomas, J., 2002. Equity Valuation Using Multiples. Journal of Accounting Research 40, 135–172. doi:10.1111/1475-679X.00042

Majd, S., Pindyck, R., 1987. Time to Build, Option Value, and Investment Decisions. Journal of Financial Economics 18: 7-27

Matsusaka J, G., Nanda, V., 1997. Internal Capital Markets and Corporate Refocusing. Journal of Financial Intermediation 11, 176-211. doi:10.1006/jfin.2001.0333

McGahan, A.M., Porter, M.E., 1997. How Much Does Industry Matter, Really? Strategic Management Journal 18, 15–30.

Mclean, R.D., Pontiff, J., 2016. Does Academic Research Destroy Stock Return Predictability? Journal of Finance 71, 5–32. doi:10.1111/jofi.12365

Miller, M, H., ROCK, K., 1985. Dividend Policy under Asymmetric Information. Journal of Finance 40, 1031-1051. doi:10.1111/j.1540-6261.1985.tb02362.x

Miller, M., Modigliani, F., 1961. Dividend Policy, Growth, and the Valuation of Shares. Journal of Business 34, 411-433.

Mishkin, F., 1983. A Rational Expectations Approach to Macroeconometrics: Testing Policy Effectiveness and Efficient Markets Models. Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Modigliani, F., Miller, M., 1958. The Cost of Capital, Corporation Finance and the Theory of Investment. American Economic Review 48, 261-297.

Mueller, D.C., Raunig, B., 1999. Heterogeneities within Industries and Structure-Performance Models. Review of Industrial Organization 15, 303–320. doi:10.1023/A:1007775731338

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28. doi:10.1016/j.jfineco.2013.01.003

Olsen, C.S., Clark, A.E., Thomas, A.M., Cook, L.J., 2012. Comparing Least-squares and Quantile Regression Approaches to Analyzing Median Hospital Charges. Academic Emergency Medicine 19, 866–875. doi:10.1111/j.1553-2712.2012.01388.x

Penman, S.H., 1991. An Evaluation of Accounting Rate-of-return. Journal of Accounting, Auditing & Finance 6, 233–255.

Penrose, E. T., 1959. The Theory of the Growth of the Firm. New York: John Wiley.

Pinnuck, M., Lillis, A, M., 2007. Profits versus losses: does reporting an accounting loss act as a heuristic trigger to exercise the abandonment option and divest employees? Accounting Review 82, 1031-1053.

Pohlman, L., Ma, L., 2010. Return forecasting by quantile regression. Journal of Investing 19, 116–121.

Pozen, R. C., 1994. Institutional investors: The reluctant activists, Harvard Business Review 71, 140–149.

Rajan, R., Servaes, H., Zingales, L., 2000. The Cost of Diversity: The Diversification Discount and Inefficient Investment. Journal of Finance 55, 35–80. doi:10.1111/0022-1082.00200

Richardson, S., Tuna, İ., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: A review of recent research advances. Journal of Accounting and Economics 50, 410–454.

Roberts, M, R., Whited, T, M., 2012. Endogeneity in Empirical Corporate Finance. Simon School Working Paper No. FR 11-29. Available at SSRN: https://ssrn.com/abstract=1748604

Roosenboom, P., 2012. Valuing and pricing IPOs. Journal of Banking & Finance 36, 1653–1664. doi:10.1016/j.jbankfin.2012.01.009

Rozeff, M, S., 1982. Growth, Beta and Agency Costs as Determinants of Dividend Payout Ratios. Journal of Financial Research 5, 249-259. doi:10.1111/j.1475-6803.1982.tb00299.x

Rumelt, R.P., 1991. How Much Does Industry Matter? Strategic Management Journal 12, 167–185.

Samuelson, P., 1948. Economics, an Introductory Analysis (McGraw- Hill, New York).

Scharfstein, D, S., Stein, J, C., 2000. The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment. Journal of Finance 55, 2537–2564. doi:10.1111/0022-1082.00299

Scharfstein, D, S., Stein, J, C., 2000. The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment. Journal of Finance 55, 2537–2564. doi:10.1111/0022-1082.00299

Schmalensee, R., 1985. Do Markets Differ Much. American Economic Association 75, 341–351.

Schröder, D., Yim, A., 2017. Industry Effects in Firm and Segment Profitability Forecasting. Contemporary Accounting Research, forthcoming.

Servaes, H., 1996. The Value of Diversification During the Conglomerate Merger Wave. Journal of Finance 51, 1201–1225. doi:10.1111/j.1540-6261.1996.tb04067.x

Shin, H, H., Stulz R, M., 1998. Are Internal Capital Markets Efficient? Quarterly Journal of Economics 113, 531–552, https://doi.org/10.1162/003355398555676

Shleifer, A., Vishny, R, W., 1989. Management entrenchment: The Case of Manager-specific Investments, Journal of Financial Economics 25, 123-139, ISSN 0304-405X, https://doi.org/10.1016/0304-405X(89)90099-8.

Shleifer, A., Vishny, R, W., 1992. Liquidation Values and Debt Capacity: A Market Equilibrium Approach. Journal of Finance 47, 1343-366. doi:10.2307/2328943.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52, 327–340.

Shumway, T., Warther, V.A., 1999. The delisting bias in CRSP's Nasdaq data and its implications for the size effect. Journal of Finance 54, 2361–2379.

Singh, M., Nejadmalayeri, A., Mathur, I., 2007. Performance Impact of Business Group Affiliation: An Analysis of the Diversification-performance Link in a Developing Economy, Journal of Business Research 60 339-347, ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2006.10.024.

Soliman, M.T., 2004. Using Industry-Adjusted DuPont Analysis to Predict Future Profitability (SSRN Scholarly Paper No. ID 456700). Social Science Research Network, Rochester, NY.

Stein J, C., 2012, Internal Capital Market and the Competition for Corporate Resources. Journal of Finance 52, 111-133. doi:10.1111/j.1540-6261.1997.tb03810.x

Stulz, R, M., 1990. Managerial discretion and optimal financing policies, Journal of Financial Economics 26, 3-27, ISSN 0304-405X, https://doi.org/10.1016/0304-405X(90)90011-N.

Subramaniam V., Tang, T.T., Yue, H., Zhou, X., 2011. Firm Structure and Corporate Cash Holdings. Journal of Corporate Finance 17, 759-773, ISSN 0929-1199. https://doi.org/10.1016/j.jcorpfin.2010.06.002.

Sundaramurthy, C., Rechner, P., Wang, W., 1996. Governance Antecedents of Board Entrenchment: The Case of Classified Board Provisions. Journal of Management 22, 783-799.

Teece, D, J., 1980. Economies of scope and the scope of the enterprise. Journal of Economic Behavior and Organization 1, 223-247.

Teece, D, J., 1982. Towards an economic theory of the multiproduct firm. Journal of Economic Behavior and Organization 3, 39-63.

Thomas, S., 2002. Firm Diversification and Asymmetric Information: Evidence from Analysts' Forecasts and Earnings Announcements. Journal of Financial Economics 64, 373-396, ISSN 0304-405X, https://doi.org/10.1016/S0304-405X(02)00129-0.

Tong, Z, X., 2011. Firm Diversification and the value of Corporate Cash Holdings. Journal of Corporate Finance 17, 741-758.

Villalonga, B., 2004. Does Diversification Cause the "Diversification Discount"?" Financial Management 33, 5-27. http://www.jstor.org/stable/3666156.

Wernerfelt, B., Montgomery, C, A., 1988. Tobin's Q and the Importance of Focus in Firm Performance. American Economic Review 78,246-50. http://www.jstor.org/stable/1814713.

Weston, J, F., 1970. The nature and significance of conglomerate firms St. John's Law Review 44, 66-80

Whited, T., 2001. Is it inefficient investment that causes the diversification discount? Journal of Finance 56, 1667-1691

Wysocki, P, D., 2001. Real Options and the Informativeness of Segment Disclosures. Working Paper, MIT, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=141044.

Yu, K., Lu, Z., Stander, J., 2003. Quantile regression: applications and current research areas. Journal of the Royal Statistical Society: Series D (The Statistician) 52, 331–350.

Zwiebel, J., 1996. Dynamic Capital Structure under Managerial Entrenchment. The American Economic Review 86, 1197-1215. Retrieved from http://www.jstor.org/stable/2118286

Appendix 3.1

Long-term profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression

	2 years ahead			3 years ahead			4 years ahead		5 years ahead		
	Value		*p-Value*	Value		*p-Value*	Value		Value		*p-Value*
GP											
Mean	0.342%	***	0.000	0.365%	***	0.000	0.361%	***	0.393%	***	0.000
Median	0.365%	***	0.000	0.379%	***	0.000	0.377%	***	0.409%	***	0.000
OP											
Mean	0.076%	***	0.000	0.054%	***	0.000	0.042%	***	0.032%	***	0.000
Median	0.066%	***	0.000	0.047%	***	0.000	0.042%	***	0.034%	***	0.000
CbOP_BS											
Mean	0.081%	***	0.000	0.075%	***	0.000	0.059%	***	0.057%	***	0.000
Median	0.088%	***	0.000	0.087%	***	0.000	0.068%	***	0.065%	***	0.000
CbOP_CF											
Mean	0.096%	***	0.000	0.087%	***	0.000	0.073%	***	0.069%	***	0.000
Median	0.101%	***	0.000	0.095%	***	0.000	0.083%	***	0.077%	***	0.000
RNOA											
Mean	0.139%	***	0.000	0.111%	***	0.000	0.094%	***	0.079%	***	0.000
Median	0.126%	***	0.000	0.112%	***	0.000	0.093%	***	0.071%	***	0.000
ROE											
Mean	0.338%	***	0.000	0.288%	***	0.000	0.267%	***	0.231%	***	0.000
Median	0.127%	***	0.000	0.106%	***	0.000	0.102%	***	0.093%	***	0.000

This table reports the 2 years to 5 years ahead profitability forecast improvements of economy-wide quantile regression (the alternative approach) over economy-wide OLS regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both forecasting approaches use the same set of predictor variables like those in Fairfield et al (2009). Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see table 1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 3.2
Profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression by using the decomposition model

	2003			2006			2017		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
Model 1									
Mean	0.045%	***	0.000	0.074%	***	0.000	0.078%	***	0.000
Median	0.069%	***	0.000	0.094%	***	0.000	0.103%	***	0.000
Model 2									
Mean	0.044%	***	0.000	0.072%	***	0.000	0.076%	***	0.000
Median	0.070%	***	0.000	0.092%	***	0.000	0.095%	***	0.000
Model 3									
Mean	0.044%	***	0.000	0.070%	***	0.000	0.072%	***	0.000
Median	0.067%	***	0.000	0.088%	***	0.000	0.088%	***	0.000

This table reports the profitability forecast improvements of economy-wide quantile regression (the alternative approach) over economy-wide OLS regression (the benchmark approach). The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. The basic model of both forecasting approaches use the same set of predictor variables like those in Sloan (2006) by decomposing ROA into ACCRUAL and CFO. Regardless of the forecasting approaches, the underlying predictor variable *PREDGSL* (i.e., the predicted growth in sales) is constructed in the same way by industry-specific OLS regression. Industries are defined using the first-digit Standard Industry Classification (SIC). Firm-specific forecasts are obtained in two steps. The basic model and the models with additional predictive variables are as below:
Model 1: ROA ~ Accrual.L1 + CFO.L1,
Model 2: ROA~ Accrual.L1 + CFO.L1 + LOSS_D.L1,
Model 3: ROA~ Accrual.L1 +CFO.L1 + LOSS_D.L1 + PREDGSL,
First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see table 1 for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 3.3
Hedged portfolio robustness tests based on negative DIFF

Panel A: Portfolio returns based on profitabiliy forecasts and forecast improvements: Economy-wide quantile versus economy-wide OLS regression

	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)			Average excess: High − Low			Carhart 4-factor alpha			Fama and French 5-factor alpha		
	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*
GP	0.585%	**	0.036	0.601%	**	0.031	-0.016%		0.940	-0.038%		0.856	-0.077%		0.706
OP	0.069%		0.850	0.379%		0.241	-0.311%	*	0.079	-0.417%	**	0.015	-0.420%	**	0.015
CbOP_BS	0.308%		0.326	0.520%	*	0.090	-0.211%		0.207	-0.232%		0.166	-0.394%	**	0.022
CbOP_CF	0.214%		0.508	0.464%		0.115	-0.250%		0.119	-0.290%	*	0.066	-0.377%	**	0.020
RNOA	0.017%		0.963	0.333%		0.250	-0.316%		0.121	-0.334%	*	0.069	-0.391%	**	0.042
ROE	-0.037%		0.926	0.276%		0.302	-0.313%		0.192	-0.448%	**	0.031	-0.391%	*	0.060

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *DIFF* variable, defined as t**he benchmark-approach forecast in excess of the alternative-approach forecast**, after confining to firms with **negative** forecast improvements. For this table, the alternative forecasting approach is economy-wide quantile regression, whereas the benchmark approach is economy-wide OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

Appendix 3.3
Hedged portfolio robustness tests based on negative DIFF

Panel B: Portfolio returns based on profitabiliy forecasts and forecast improvements: Industry-specific quantile versus industry-specific OLS regression

	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)			Average excess: High − Low		Carhart 4-factor alpha		Fama and French 5-factor alpha		
	Return		*p-Value*	Return		*p-Value*	Return	*p-Value*	Return	*p-Value*	Return		*p-Value*
GP	0.516%	*	0.071	0.546%	**	0.046	-0.030%	0.888	-0.058%	0.779	-0.021%		0.917
OP	0.080%		0.808	0.342%		0.326	-0.262%	0.177	-0.321%	0.108	-0.231%		0.257
CbOP_BS	0.292%		0.335	0.462%		0.104	-0.170%	0.295	-0.194%	0.224	-0.284%	*	0.084
CbOP_CF	0.469%		0.119	0.480%	*	0.083	-0.011%	0.947	-0.103%	0.536	-0.130%		0.443
RNOA	0.313%		0.306	0.276%		0.391	0.037%	0.849	-0.014%	0.940	-0.089%		0.644
ROE	-0.001%		0.998	0.118%		0.727	-0.119%	0.599	-0.230%	0.270	-0.184%		0.396

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *-DIFF* variable, defined as **the benchmark-approach forecast in excess of the laternative-approach forecast**, after confining to firms with **negative** forecast improvements. For this table, the alternative forecasting approach is industry-specific quantile regression, whereas the benchmark approach is industry-specific OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

Appendix 3.4

Hedged portfolio robustness tests: full sample

Panel A: Portfolio returns based on profitabiliy forecasts and forecast improvements: Economy-wide quantile versus economy-wide OLS regression full sample															
	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)			Average excess: High − Low			Carhart 4-factor alpha			Fama and French 5-factor alpha		
	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*
GP	0.767%	***	0.002	0.334%		0.195	0.433%	***	0.009	0.408%	***	0.004	0.420%	***	0.003
OP	0.870%	***	0.000	0.164%		0.635	0.706%	***	0.000	0.678%	***	0.000	0.678%	***	0.000
CbOP_BS	0.802%	***	0.002	0.267%		0.354	0.535%	***	0.000	0.553%	***	0.000	0.665%	***	0.000
CbOP_CF	0.846%	***	0.001	0.218%		0.447	0.627%	***	0.000	0.628%	***	0.000	0.676%	***	0.000
RNOA	0.804%	***	0.001	0.185%		0.583	0.619%	***	0.003	0.596%	***	0.000	0.601%	***	0.000
ROE	0.809%	***	0.000	0.054%		0.876	0.755%	***	0.000	0.761%	***	0.000	0.755%	***	0.000

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *DIFF* variable, defined as the alternative-approach forecast in excess of the benchmark-approach forecast, **without** confining to firms with positive forecast improvements. For this table, the alternative forecasting approach is economy-wide quantile regression, whereas the benchmark approach is economy-wide OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

Appendix 3.4

Hedged portfolio robustness tests: full sample

Panel B: Portfolio returns based on profitabiliy forecasts and forecast improvements: Industry-specific quantile versus industry-specific OLS regression

	Long portfolio: High (quartiles)			Short portfolio: Low (quartiles)		Average excess: High − Low			Carhart 4-factor alpha			Fama and French 5-factor alpha		
	Return		*p-Value*	Return	*p-Value*	Return		*p-Value*	Return		*p-Value*	Return		*p-Value*
GP	0.774%	***	0.002	0.452%	0.116	0.321%	*	0.056	0.342%	**	0.016	0.387%	***	0.009
OP	0.837%	***	0.001	0.251%	0.432	0.586%	***	0.001	0.526%	***	0.001	0.536%	***	0.001
CbOP_BS	0.805%	***	0.001	0.260%	0.379	0.546%	***	0.001	0.551%	***	0.000	0.593%	***	0.000
CbOP_CF	0.840%	***	0.000	0.239%	0.427	0.601%	***	0.000	0.592%	***	0.000	0.603%	***	0.000
RNOA	0.796%	***	0.001	0.341%	0.230	0.454%	***	0.005	0.404%	***	0.007	0.395%	**	0.011
ROE	0.734%	***	0.002	0.292%	0.253	0.443%	***	0.005	0.448%	***	0.002	0.513%	***	0.000

This table reports value-weighted excess returns, Carhart (1997) 4-factor alphas, and Fama and French (2015) 5-factor alphas for portfolios sorted by the *DIFF* variable, defined as the alternative-approach forecast in excess of the benchmark-approach forecast, **without** confining to firms with positive forecast improvements. For this table, the alternative forecasting approach is industry-specific quantile regression, whereas the benchmark approach is industry-specific OLS regression. Firms with positive forecast improvements are those with the absolute forecast error from the benchmark approach larger than that from the alternative approach. At the end of each June, we sort stocks of the firms with positive forecast improvements into quartiles and hold the portfolio for the following year. The sample starts in July 1989 and ends in June 2016.

Appendix 3.5
Sample selection and descriptive statistics, 1985-2016

Panel A: Descriptive statistics for quarterly forecasts

	Q1					Q2				
Variable	N	Mean	Median	Std. Dev.	Skewness	N	Mean	Median	Std. Dev.	Skewness
IBES_EPS_P	*46846*	0.9%	1.1%	0.03	-1.90	*44960*	1.2%	1.4%	0.03	-1.02
Analysts Mean Consensus	*46846*	0.9%	1.1%	0.03	-2.62	*44960*	1.3%	1.4%	0.03	-0.66
Analysts Median Consensus	*46846*	0.9%	1.1%	0.03	-2.85	*44960*	1.3%	1.4%	0.03	-0.74
IBES_EPS_CH_SEQ	*49189*	2.3%	2.3%	0.08	-1.07	*46831*	3.1%	2.8%	0.08	-0.61
Analysts Mean Consensus	*49189*	2.3%	2.2%	0.07	-1.03	*46831*	3.5%	2.9%	0.07	0.16
Analysts Median Consensus	*49189*	2.3%	2.2%	0.07	-1.11	*46831*	3.5%	2.9%	0.07	0.09

This panel gives an overview of the data used to compute the forecast improvements for the period from 1985 to 2016. For quarterly forecasts, IBES_EPS_P is the actual quarterly earnings per share (EPS) from I/B/E/S deflated by stock price. Stock price is lagged two years' fiscal annual closing price collected from I/B/E/S. IBES_EPS_CH_SEQ is the actual quarterly earnings per share (EPS) from I/B/E/S multiple by number of shares outstanding in the previous year deflated by average of book value of equity of the previous one year and two years. Similarly, Analysts Mean (Median) Consensus is the mean (median) of analysts' forecasts of EPS and deflate the same deflators which is consistent with the actual deflated quarterly earnings.

Appendix 3.5 (continued)
Sample selection and descriptive statistics, 1985-2016

Panel A: Descriptive statistics for quarterly forecasts

	Q3					Q4				
Variable	N	Mean	Median	Std. Dev.	Skewness	N	Mean	Median	Std. Dev.	Skewness
IBES_EPS_P	*44843*	1.6%	1.4%	0.03	-0.61	*46021*	1.2%	1.4%	0.04	-1.59
Analysts Mean Consensus	*44843*	1.6%	1.6%	0.02	1.42	*46021*	1.8%	1.7%	0.03	1.98
Analysts Median Consensus	*44843*	1.3%	1.6%	0.02	1.45	*46021*	1.8%	1.7%	0.03	2.03
IBES_EPS_CH_SEQ	*46552*	3.2%	2.9%	0.09	-0.55	*47753*	3.0%	2.9%	0.10	-1.03
Analysts Mean Consensus	*46552*	4.2%	3.4%	0.07	0.52	*47753*	4.8%	3.6%	0.08	1.17
Analysts Median Consensus	*46552*	4.2%	3.4%	0.07	0.53	*47753*	4.8%	3.6%	0.08	1.12

This panel gives an overview of the data used to compute the forecast improvements for the period from 1985 to 2016. For quarterly forecasts, IBES_EPS_P is the actual quarterly earnings per share (EPS) from I/B/E/S deflated by stock price. Stock price is lagged two years' fiscal annual closing price collected from I/B/E/S. IBES_EPS_CH_SEQ is the actual quarterly earnings per share (EPS) from I/B/E/S multiple by number of shares outstanding in the previous year deflated by average of book value of equity of the previous one year and two years. Similarly, Analysts Mean (Median) Consensus is the mean (median) of analysts' forecasts of EPS and deflate the same deflators which is consistent with the actual deflated quarterly earnings.

Appendix 3.5

Panel B: Descriptive statistics for long-term forecasts

	Y1					Y2					Y3				
Variable	N	Mean	Median	Std. Dev.	Skewness	N	Mean	Median	Std. Dev.	Skewness	N	Mean	Median	Std. Dev.	Skewness
IBES_EPS_P	*52190*	4.2%	5.4%	0.10	-1.11	*46692*	5.0%	5.7%	0.11	-0.81	*16384*	5.8%	6.4%	0.11	-0.25
Analysts Mean Consensus	*52190*	5.3%	6.0%	0.08	-1.24	*46692*	6.5%	7.1%	0.07	-2.83	*16384*	7.7%	7.2%	0.09	-0.33
Analysts Median Consensus	*52190*	5.3%	6.0%	0.08	-1.23	*46692*	6.5%	7.0%	0.07	-2.79	*16384*	7.7%	7.2%	0.09	-0.28
IBES_EPS_CH_SEQ	*53222*	10.6%	10.4%	0.22	-0.55	*43732*	11.4%	10.8%	0.21	-0.46	*15031*	14.5%	13.1%	0.22	-0.31
Analysts Mean Consensus	*53222*	13.5%	11.7%	0.20	-0.19	*43732*	16.5%	14.0%	0.18	0.11	*15031*	18.7%	16.1%	0.17	0.14
Analysts Median Consensus	*53222*	13.5%	11.7%	0.20	-0.20	*43732*	16.4%	13.9%	0.18	0.10	*15031*	18.6%	16.1%	0.17	0.10

This panel gives an overview of the data used to compute the forecast improvements for the period from 1985 to 2016. For annual forecasts, IBES_EPS_P is the actual annual earnings per share (EPS) from I/B/E/S deflated by stock price. Stock price is lagged two years' fiscal annual closing price collected from I/B/E/S. IBES_EPS_CH_SEQ is the actual quarterly earnings per share (EPS) from I/B/E/S multiple by number of shares outstanding in the previous year deflated by average of book value of equity of the previous one year and two years. Similarly, Analysts Mean (Median) Consensus is the mean (median) of analysts' forecasts of EPS and deflate the same deflators which is consistent with the actual deflated annual earnings.

Appendix 3.5 (continued)

Panel B: Descriptive statistics for long-term forecasts

	Y4					Y5				
Variable	N	Mean	Median	Std. Dev.	Skewness	N	Mean	Median	Std. Dev.	Skewness
IBES_EPS_P	*5638*	6.5%	7.2%	0.12	-0.50	*3021*	7.1%	7.6%	0.12	-0.02
Analysts Mean Consensus	*5638*	8.8%	7.7%	0.10	1.04	*3021*	10.2%	8.0%	0.11	1.90
Analysts Median Consensus	*5638*	8.7%	7.7%	0.10	0.71	*3021*	10.1%	8.0%	0.11	1.69
IBES_EPS_CH_SEQ	*5070*	15.5%	13.9%	0.22	-0.49	*2721*	16.1%	14.1%	0.22	-0.26
Analysts Mean Consensus	*5070*	19.3%	15.9%	0.17	0.46	*2721*	19.8%	15.7%	0.17	1.16
Analysts Median Consensus	*5070*	19.3%	15.9%	0.17	0.44	*2721*	19.7%	15.6%	0.17	1.07

This panel gives an overview of the data used to compute the forecast improvements for the period from 1985 to 2016. For annual forecasts, IBES_EPS_P is the actual quarterly earnings per share (EPS) from I/B/E/S deflated by stock price. Stock price is lagged two years' fiscal annual closing price collected from I/B/E/S. IBES_EPS_CH_SEQ is the actual quarterly earnings per share (EPS) from I/B/E/S multiple by number of shares outstanding in the previous year deflated by average of book value of equity of the previous one year and two years. Similarly, Analysts Mean (Median) Consensus is the mean (median) of analysts' forecasts of EPS and deflate the same deflators which is consistent with the actual deflated annual earnings.

Appendix 3.6
Profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression: IBES-based ROE by Method 2

Panel A: Quarterly forecasting

Q1	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-1.504%	***	0.000	-1.515%	***	0.000	0.046%	***	0.000
Median	-0.820%	***	0.000	-0.826%	***	0.000	0.040%	***	0.000
PRED_OLS									
Mean	-1.551%	***	0.000	-1.562%	***	0.000			
Median	-0.890%	***	0.000	-0.895%	***	0.000			

Q2	Analysts’ mean consensus			Analysts’ median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-1.161%	***	0.000	-1.155%	***	0.000	0.035%	***	0.000
Median	-0.612%	***	0.000	-0.610%	***	0.000	0.032%	***	0.000
PRED_OLS									
Mean	-1.196%	***	0.000	-1.189%	***	0.000			
Median	-0.663%	***	0.000	-0.661%	***	0.000			

Q3	Analysts’ mean consensus			Analysts’ median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.856%	***	0.000	-0.858%	***	0.000	0.041%	***	0.000
Median	-0.473%	***	0.000	-0.476%	***	0.000	0.039%	***	0.000
PRED_OLS									
Mean	-0.898%	***	0.000	-0.899%	***	0.000			
Median	-0.537%	***	0.000	-0.537%	***	0.000			

Q4	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-0.597%	***	0.000	-0.604%	***	0.000	0.090%	***	0.000
Median	-0.392%	***	0.000	-0.394%	***	0.000	0.094%	***	0.000
PRED_OLS									
Mean	-0.688%	***	0.000	-0.694%	***	0.000			
Median	-0.513%	***	0.000	-0.515%	***	0.000			

This table reports the short-term quarterly profitability forecast improvements of economy-wide model-based forecast model (the alternative approach by row) over analysts’ consensus forecasts (the benchmark approach by column). See the main texts for the detailed description of the IBES-based ROE by Method 1 and Method 2. The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches which equals the AFE by using the column approach minus the AFE by using the row approach. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both model-based forecasting approaches use the same forecasting steps like those in Fairfield et al (2009) but simply based on a parsimonious AR (1) model. Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see main texts for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 3.6
Profitability forecast improvements of economy-wide quantile regression over economy-wide OLS regression: IBES-based ROE by Method 2

Panel B: Long-term forecasting

Y1	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	-2.254%	***	0.000	-2.260%	***	0.000	0.367%	***	0.000
Median	-1.285%	***	0.000	-1.285%	***	0.000	0.393%	***	0.000
PRED_OLS									
Mean	-2.621%	***	0.000	-2.627%	***	0.000			
Median	-1.792%	***	0.000	-1.793%	***	0.000			

Y2	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	1.405%	***	0.000	1.406%	***	0.000	0.305%	***	0.000
Median	0.545%	***	0.000	0.546%	***	0.000	0.340%	***	0.000
PRED_OLS									
Mean	1.100%	***	0.000	1.101%	***	0.000			
Median	0.263%	***	0.000	0.265%	***	0.000			

Y3	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	1.913%	***	0.000	1.894%	***	0.000	0.436%	***	0.000
Median	0.813%	***	0.000	0.798%	***	0.000	0.433%	***	0.000
PRED_OLS									
Mean	1.477%	***	0.000	1.458%	***	0.000			
Median	0.476%	***	0.000	0.468%	***	0.000			

Y4	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	2.313%	***	0.000	2.310%	***	0.000	0.388%	***	0.000

Median	0.855%	***	0.000	0.855%	***	0.000	0.365%	***	0.000
PRED_OLS									
Mean	1.926%	***	0.000	1.922%	***	0.000			
Median	0.585%	***	0.000	0.586%	***	0.000			

Y5	Analysts' mean consensus			Analysts' median consensus			PRED_OLS		
	Value		*p-Value*	Value		*p-Value*	Value		*p-Value*
PRED_QR									
Mean	2.470%	***	0.000	2.469%	***	0.000	2.113%	***	0.000
Median	0.924%	***	0.000	0.952%	***	0.000	0.323%	***	0.000
PRED_OLS									
Mean	2.112%	***	0.000	0.357%	***	0.000			
Median	0.732%	***	0.000	0.764%	***	0.000			

This table reports the long-term annual profitability forecast improvements of economy-wide model-based forecast model (the alternative approach by row) over analysts' consensus forecasts (the benchmark approach by column). See the main texts for the detailed description of the IBES-based ROE by Method 1 and Method 2. The forecast improvement (FI) is measured through a matched-pair comparison of the absolute forecast errors (AFE) from the two competing approaches which equals the AFE by using the column approach minus the AFE by using the row approach. A positive FI means the AFE from the benchmark approach is larger than that from the alternative approach. Both model-based forecasting approaches use the same forecasting steps like those in Fairfield et al (2009) but simply based on a parsimonious AR (1) model. Firm-specific forecasts are obtained in two steps. First, the coefficients of a forecasting model are estimated for each year from 1989 to 2016 on a rolling basis using data of the previous 10 years. Next, the estimated coefficients for a year are applied on a firm's data of the previous year to obtain a firm-specific forecast for the current year. The mean and median FIs of all firm-years in the sample are reported for different profitability measures (see main texts for the definitions of the measures). The test on the mean FI is a regression-based t-test using robust standard errors controlling for two-way clustering by firm and year. The test on the median FI is the Wilcoxon signed rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.1

Loss reversal and firm structure: logistic model under Second- Stage of Heckman Test

Panel A: Loss reversal and firm structure full sample

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.415	0.139	0.003	***	0.426	0.158	0.007	***	0.456	0.158	0.004	***
ROA	4.086	0.266	0.000	***	4.084	0.265	0.000	***	4.084	0.265	0.000	***
PAST_ROA	0.191	0.101	0.057	*	0.197	0.101	0.051	*	0.196	0.101	0.052	*
SIZE	-0.029	0.010	0.005	***	-0.030	0.010	0.003	***	-0.030	0.010	0.003	***
SALESGROWTH	0.014	0.034	0.674		0.014	0.034	0.687		0.014	0.034	0.689	
FIRSTLOSS	-0.024	0.049	0.629		-0.024	0.049	0.625		-0.024	0.049	0.629	
LOSS_SEQ	-0.159	0.015	0.000	***	-0.158	0.015	0.000	***	-0.158	0.015	0.000	***
DIVDUM	0.165	0.038	0.000	***	0.157	0.038	0.000	***	0.156	0.038	0.000	***
LEVERAGE	0.175	0.076	0.022	**	0.160	0.076	0.036	**	0.158	0.076	0.038	**
T_Q	0.000	0.010	0.967		0.001	0.010	0.954		0.001	0.010	0.941	
CAPX	-0.707	0.224	0.002	***	-0.687	0.223	0.002	***	-0.686	0.223	0.002	***
SPECIAL_ITEM	-6.051	0.325	0.000	***	-6.056	0.325	0.000	***	-6.057	0.325	0.000	***
CREDIT_RATING	0.250	0.033	0.000	***	0.246	0.033	0.000	***	0.245	0.033	0.000	***
DIVERSITY_D	-0.044	0.046	0.334									
HERF_ASSET					-0.047	0.098	0.634					
HERF_SALES									-0.082	0.098	0.404	
AB_D	-1.241	0.044	0.000	***	-0.732	0.145	0.000	***	-0.768	0.147	0.000	***
AB_D*DIVERSITY_D	0.211	0.078	0.007	***								
AB_D*HERF_ASSET					-0.516	0.162	0.001	***				
AB_D*HERF_SALES									-0.473	0.164	0.004	***
INVERSE MILLS RATIO	-0.235	0.072	0.001	***	-0.208	0.072	0.004	***	-0.206	0.072	0.004	***
Diversified & AB_D=1												
marginal effects	0.020	0.007	0.003	***	-0.06	0.012	0.000	***	-0.056	0.012	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression under the second stage of Heckman test. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 34820. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.1

Loss reversal and firm structure: logistic model under Second- Stage of Heckman Test

Panel B: Loss reversal and firm structure under over-investment problem (High agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.111	0.223	0.620		0.017	0.247	0.944		0.082	0.246	0.738	
ROA	5.450	0.407	0.000	***	5.430	0.406	0.000	***	5.425	0.405	0.000	***
PAST_ROA	0.415	0.170	0.014	**	0.418	0.170	0.014	**	0.418	0.170	0.014	**
SIZE	-0.040	0.016	0.010	**	-0.042	0.016	0.007	***	-0.042	0.016	0.007	***
SALESGROWTH	-0.021	0.059	0.720		-0.023	0.059	0.702		-0.023	0.059	0.697	
FIRSTLOSS	0.009	0.068	0.890		0.008	0.068	0.910		0.008	0.068	0.905	
LOSS_SEQ	-0.154	0.022	0.000	***	-0.154	0.022	0.000	***	-0.154	0.022	0.000	***
DIVDUM	0.141	0.054	0.009	***	0.133	0.054	0.013	**	0.131	0.054	0.015	**
LEVERAGE	0.331	0.143	0.020	**	0.321	0.143	0.024	**	0.319	0.142	0.025	**
T_Q	0.177	0.093	0.058	*	0.179	0.093	0.055	*	0.180	0.093	0.052	*
CAPX	-1.636	0.453	0.000	***	-1.607	0.452	0.000	***	-1.603	0.451	0.000	***
SPECIAL_ITEM	-7.102	0.496	0.000	***	-7.095	0.496	0.000	***	-7.092	0.496	0.000	***
CREDIT_RATING	0.238	0.047	0.000	***	0.236	0.047	0.000	***	0.235	0.047	0.000	***
DIVERSITY_D	-0.086	0.065	0.186									
HERF_ASSET					0.056	0.138	0.685					
HERF_SALES									-0.025	0.137	0.856	
AB_D	-1.116	0.060	0.000	***	-0.794	0.193	0.000	***	-0.876	0.194	0.000	***
AB_D*DIVERSITY_D	0.104	0.105	0.321									
AB_D*HERF_ASSET					-0.337	0.217	0.121					
AB_D*HERF_SALES									-0.241	0.218	0.268	
INVERSE MILLS RATIO	-0.046	0.108	0.668		-0.022	0.107	0.837		-0.016	0.107	0.880	
Div&AB_D=1 marginal effect	0.0004	0.011	0.969		-0.033	0.022	0.138		-0.024	0.023	0.282	

This table presents the results of loss reversal probability using the annual estimation of logistic regression under the second stage of Heckman test with the firms having agency problems. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herdial Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of emplyees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 16523. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.1

Loss reversal and firm structure: logistic model under Second- Stage of Heckman Test

Panel C: Loss reversal and firm structure under no over-investment problem (no agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.362	0.191	0.059	*	0.447	0.217	0.040	**	0.452	0.218	0.038	**
ROA	3.447	0.322	0.000	***	3.455	0.322	0.000	***	3.454	0.322	0.000	***
PAST_ROA	0.120	0.128	0.347		0.128	0.128	0.319		0.126	0.128	0.324	
SIZE	-0.021	0.014	0.131		-0.022	0.014	0.106		-0.023	0.014	0.101	
SALESGROWTH	-0.002	0.042	0.969		-0.002	0.042	0.960		-0.002	0.042	0.960	
FIRSTLOSS	-0.090	0.071	0.208		-0.088	0.071	0.218		-0.088	0.071	0.215	
LOSS_SEQ	-0.161	0.022	0.000	***	-0.160	0.022	0.000	***	-0.160	0.022	0.000	***
DIVDUM	0.184	0.054	0.001	***	0.173	0.054	0.001	***	0.175	0.054	0.001	***
LEVERAGE	0.558	0.129	0.000	***	0.543	0.129	0.000	***	0.543	0.129	0.000	***
T_Q	-0.010	0.012	0.402		-0.010	0.012	0.408		-0.009	0.012	0.413	
CAPX	-0.569	0.264	0.031	**	-0.555	0.264	0.035	**	-0.556	0.264	0.035	**
SPECIAL_ITEM	-5.827	0.426	0.000	***	-5.837	0.426	0.000	***	-5.834	0.426	0.000	***
CREDIT_RATING	0.248	0.048	0.000	***	0.243	0.048	0.000	***	0.242	0.048	0.000	***
DIVERSITY_D	-0.016	0.065	0.808									
HERF_ASSET					-0.126	0.142	0.375					
HERF_SALES									-0.127	0.143	0.376	
AB_D	-1.404	0.069	0.000	***	-0.568	0.227	0.012	**	-0.548	0.231	0.018	**
AB_D*DIVERSITY_D	0.392	0.121	0.001	***								
AB_D*HERF_ASSET					-0.829	0.254	0.001	***				
AB_D*HERF_SALES									-0.848	0.257	0.001	***
INVERSE MILLS RATIO	-0.289	0.103	0.005	***	-0.257	0.104	0.013	**	-0.259	0.103	0.012	**
Diversified & AB_D=1 marginal effects	0.039	0.009	0.000	***	-0.081	0.014	0.000	***	-0.081	0.014	0.000	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression under the second stage of Heckman test with the firms having no potential agency problems. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herdial Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of emplyees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 18298. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.2

Loss reversal and firm structure: logistic model and agency problem

Panel A: Loss reversal and firm structure among low payout firms (high agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.012	0.307	0.969		0.057	0.336	0.865		0.053	0.336	0.874	
ROA	5.079	0.829	0.000	***	5.057	0.826	0.000	***	5.061	0.826	0.000	***
PAST_ROA	0.029	0.281	0.919		0.028	0.281	0.921		0.028	0.281	0.919	
SIZE	0.022	0.019	0.241		0.019	0.019	0.329		0.019	0.019	0.324	
SALESGROWTH	-0.009	0.090	0.920		-0.009	0.090	0.921		-0.008	0.090	0.929	
FIRSTLOSS	0.099	0.104	0.342		0.099	0.104	0.341		0.099	0.104	0.342	
LOSS_SEQ	-0.174	0.038	0.000	***	-0.172	0.038	0.000	***	-0.173	0.038	0.000	***
DIVDUM	0.416	0.247	0.093	*	0.422	0.247	0.088	*	0.420	0.247	0.089	*
LEVERAGE	-0.548	0.165	0.001	***	-0.564	0.165	0.001	***	-0.564	0.165	0.001	***
T_Q	-0.058	0.029	0.043	**	-0.057	0.029	0.044	**	-0.057	0.029	0.044	**
CAPX	-1.105	0.460	0.016	***	-1.050	0.459	0.022	**	-1.063	0.459	0.021	**
SPECIAL_ITEM	-8.275	0.872	0.000	***	-8.260	0.870	0.000	***	-8.263	0.871	0.000	***
CREDIT_RATING	0.256	0.070	0.000	***	0.253	0.070	0.000	***	0.254	0.070	0.000	***
DIVERSITY_D	-0.005	0.088	0.952									
HERF_ASSET					-0.123	0.173	0.476					
HERF_SALES									-0.107	0.173	0.536	
AB_D	-1.181	0.097	0.000	***	-0.907	0.238	0.000	***	-0.899	0.240	0.000	***
AB_D*DIVERSITY_D	0.153	0.145	0.291									
AB_D*HERF_ASSET					-0.262	0.282	0.353					
AB_D*HERF_SALES									-0.271	0.283	0.339	
Diversified & AB_D=1 marginal effects	0.003	0.014	0.921		-0.041	0.023	0.213		-0.038	0.024	0.386	

This table presents the results of loss reversal probability with agency problems using the annual estimation of logistic regression. Low payout is used as the proxy of agency problem which is defined as if the firm's payout ratio is below the industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Independent variables follow Joos and Plesko (2005) with additional firm level control variables and definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 7585. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.2

Loss reversal and firm structure: logistic model and agency problem

Panel B: Loss reversal and firm structure among high payout firms (no agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.051	0.091	0.574		0.082	0.149	0.583		0.149	0.150	0.319	
ROA	4.125	0.260	0.000	***	4.113	0.260	0.000	***	4.111	0.259	0.000	***
PAST_ROA	0.248	0.102	0.015	**	0.247	0.102	0.015	**	0.246	0.102	0.016	**
SIZE	-0.028	0.010	0.005	***	-0.030	0.010	0.002	***	-0.031	0.010	0.002	***
SALESGROWTH	0.002	0.035	0.952		0.002	0.035	0.952		0.002	0.035	0.960	
FIRSTLOSS	-0.062	0.053	0.249		-0.063	0.053	0.239		-0.063	0.053	0.238	
LOSS_SEQ	-0.155	0.016	0.000	***	-0.155	0.016	0.000	***	-0.155	0.016	0.000	***
DIVDUM	0.095	0.064	0.139		0.080	0.064	0.212		0.079	0.064	0.222	
LEVERAGE	0.367	0.080	0.000	***	0.348	0.080	0.000	***	0.345	0.080	0.000	***
T_Q	0.015	0.010	0.128	***	0.015	0.010	0.130		0.015	0.010	0.124	***
CAPX	-0.679	0.240	0.005	***	-0.669	0.240	0.005	***	-0.666	0.240	0.005	***
SPECIAL_ITEM	-5.813	0.337	0.000	***	-5.812	0.337	0.000	***	-5.813	0.337	0.000	***
CREDIT_RATING	0.207	0.037	0.000	***	0.205	0.037	0.000	***	0.204	0.037	0.000	***
DIVERSITY_D	-0.010	0.051	0.841									
HERF_ASSET					-0.075	0.113	0.502					
HERF_SALES									-0.137	0.113	0.226	
AB_D	-1.257	0.048	0.000	***	-0.610	0.174	0.000	***	-0.642	0.177	0.000	***
AB_D*DIVERSITY_D	0.246	0.090	0.006	***								
AB_D*HERF_ASSET					-0.660	0.192	0.001	***				
AB_D*HERF_SALES									-0.621	0.194	0.001	***
Diversified & AB_D=1 marginal effects	0.042	0.010	0.000	***	-0.079	0.012	0.000	***	-0.079	0.012	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. Low payout is used as the proxy of agency problem which is defined as if the firm's payout ratio is below the industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variables definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 31925. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.3

Loss reversal and firm structure: logistic model and agency problem

Panel A: Loss reversal (alternative asset-based abandonment option) and firm structure under over-investment (high agency problem)	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.027	0.102	0.791		-0.004	0.144	0.975		0.086	0.144	0.549	
ROA	5.197	0.299	0.000	***	5.172	0.299	0.000	***	5.169	0.298	0.000	***
PAST_ROA	0.425	0.153	0.005	***	0.421	0.153	0.006	***	0.417	0.153	0.006	***
SIZE	-0.026	0.010	0.011	**	-0.028	0.010	0.005	***	-0.030	0.010	0.003	***
SALESGROWTH	-0.037	0.046	0.411		-0.038	0.046	0.405		-0.039	0.046	0.396	
FIRSTLOSS	0.014	0.052	0.781		0.013	0.052	0.801		0.012	0.052	0.817	
LOSS_SEQ	-0.132	0.017	0.000	***	-0.131	0.017	0.000	***	-0.131	0.017	0.000	***
DIVDUM	0.147	0.040	0.000	***	0.139	0.040	0.001	***	0.136	0.041	0.001	***
LEVERAGE	0.392	0.106	0.000	***	0.376	0.106	0.000	***	0.375	0.105	0.000	***
T_Q	0.244	0.070	0.000	***	0.246	0.070	0.000	***	0.249	0.070	0.000	***
CAPX	-1.659	0.327	0.000	***	-1.635	0.326	0.000	***	-1.633	0.326	0.000	***
SPECIAL_ITEM	-6.979	0.378	0.000	***	-6.968	0.378	0.000	***	-6.968	0.378	0.000	***
CREDIT_RATING	0.234	0.035	0.000	***	0.233	0.035	0.000	***	0.232	0.035	0.000	***
DIVERSITY_D	-0.074	0.052	0.155									
HERF_ASSET					0.031	0.109	0.776					
HERF_SALES									-0.067	0.109	0.542	
AB_D	-1.202	0.043	0.000	***	-0.857	0.139	0.000	***	-0.941	0.141	0.000	***
AB_D*DIVERSITY_D	0.148	0.076	0.050	**								
AB_D*HERF_ASSET					-0.348	0.156	0.026	**				
AB_D*HERF_SALES									-0.250	0.158	0.113	
Diversified & AB_D=1 marginal effects	0.009	0.010	0.403		-0.044	0.021	0.032	**	-0.042	0.021	0.048	**

This table presents the results of loss reversal probability with high agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Independent variables follow Joos and Plesko (2005) with additional firm level control variables and definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both total assets and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 18848. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.3

Loss reversal and firm structure: logistic model and agency problem

Panel B: Loss reversal (alternative asset-based abandonment option) and firm structure under no over-investment (no agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.018	0.080	0.827		0.185	0.134	0.166		0.188	0.134	0.160	
ROA	3.784	0.234	0.000	***	3.776	0.233	0.000	***	3.776	0.233	0.000	***
PAST_ROA	0.198	0.112	0.076	*	0.197	0.111	0.077	*	0.194	0.111	0.081	*
SIZE	0.001	0.010	0.928		-0.002	0.010	0.854		-0.002	0.010	0.860	
SALESGROWTH	0.001	0.032	0.965		0.001	0.032	0.963		0.001	0.032	0.974	
FIRSTLOSS	-0.089	0.053	0.094	*	-0.087	0.053	0.102		-0.088	0.053	0.098	*
LOSS_SEQ	-0.164	0.017	0.000	***	-0.163	0.017	0.000	***	-0.163	0.017	0.000	***
DIVDUM	0.216	0.040	0.000	***	0.204	0.040	0.000	***	0.205	0.040	0.000	***
LEVERAGE	0.733	0.094	0.000	***	0.719	0.094	0.000	***	0.725	0.094	0.000	***
T_Q	-0.011	0.009	0.213		-0.011	0.009	0.226		-0.010	0.009	0.234	
CAPX	-0.716	0.194	0.000	***	-0.696	0.194	0.000	***	-0.702	0.194	0.000	***
SPECIAL_ITEM	-6.168	0.315	0.000	***	-6.168	0.316	0.000	***	0.244	0.036	0.000	***
CREDIT_RATING	0.249	0.036	0.000	***	0.245	0.036	0.000	***	0.232	0.035	0.000	***
DIVERSITY_D	-0.012	0.050	0.818									
HERF_ASSET					-0.170	0.111	0.125					
HERF_SALES									-0.176	0.111	0.114	
AB_D	-1.591	0.048	0.000	***	-0.675	0.157	0.000	***	-0.686	0.159	0.000	***
AB_D*DIVERSITY_D	0.508	0.084	0.000	***								
AB_D*HERF_ASSET					-0.881	0.175	0.000	***				
AB_D*HERF_SALES									-0.863	0.177	0.000	***
Diversified & AB_D=1 marginal effects	0.042	0.009	0.000	***	-0.080	0.013	0.000	***	-0.077	0.014	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Independent variables follow Joos and Plesko (2005) with additional firm level control variables and definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both total assets and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 20662. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.4

Loss reversal and firm structure: logistic model and agency problem

Panel A: Loss reversal (alternative abandonment option using GSL forecasting model) and firm structure under over-investment (high agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.398	0.097	0.000		-0.366	0.127	0.004		-0.300	0.127	0.018	**
ROA	5.716	0.299	0.000	***	5.693	0.299	0.000	***	5.684	0.298	0.000	***
PAST_ROA	0.291	0.142	0.040	**	0.288	0.142	0.043	**	0.286	0.142	0.044	**
SIZE	-0.036	0.010	0.000	***	-0.039	0.010	0.000	***	-0.040	0.010	0.000	***
SALESGROWTH	-0.014	0.041	0.743		-0.014	0.041	0.727		-0.015	0.041	0.724	
FIRSTLOSS	0.043	0.049	0.384		0.042	0.049	0.393		0.042	0.049	0.397	
LOSS_SEQ	-0.111	0.016	0.000	***	-0.111	0.016	0.000	***	-0.111	0.016	0.000	***
DIVDUM	0.146	0.039	0.000	***	0.138	0.039	0.000	***	0.135	0.039	0.000	***
LEVERAGE	0.189	0.101	0.062	*	0.174	0.101	0.086	*	0.169	0.101	0.095	*
T_Q	0.389	0.068	0.000	***	0.390	0.068	0.000	***	0.393	0.068	0.000	***
CAPX	-1.086	0.320	0.001	***	-1.070	0.320	0.001	***	-1.063	0.319	0.001	***
SPECIAL_ITEM	-6.760	0.368	0.000	***	-6.750	0.368	0.000	***	-6.746	0.368	0.000	***
CREDIT_RATING	0.176	0.033	0.000	***	0.175	0.033	0.000	***	0.175	0.033	0.000	***
DIVERSITY_D	-0.041	0.039	0.294									
HERF_ASSET					-0.032	0.082	0.693					
HERF_SALES									-0.100	0.083	0.227	
AB_D	-0.043	0.076	0.567		0.332	0.215	0.122		0.270	0.216	0.212	
AB_D*DIVERSITY_D	0.130	0.124	0.296									
AB_D*HERF_ASSET					-0.394	0.247	0.111					
AB_D*HERF_SALES									-0.321	0.248	0.195	
Marginal effects	0.003	0.012	0.23		-0.008	0.023	0.127		-0.078	0.023	0.081	*

This table presents the results of loss reversal probability with high agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Independent variables follow Joos and Plesko (2005) with additional firm level control variables and definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if model-predicted sales growth decreases comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 18848. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.4

Loss reversal and firm structure: logistic model and agency problem

Panel B: Loss reversal (alternative abandonment option using GSL forecasting model) and firm structure under no over-investment (no agency problem)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.311	0.076	0.000		-0.013	0.116	0.914		-0.013	0.116	0.914	
ROA	4.047	0.231	0.000	***	4.040	0.230	0.000	***	4.040	0.230	0.000	***
PAST_ROA	0.120	0.104	0.248		0.120	0.104	0.249		0.120	0.104	0.249	
SIZE	0.004	0.009	0.667		0.002	0.009	0.831		0.002	0.009	0.831	
SALESGROWTH	0.014	0.030	0.633		0.015	0.030	0.630		0.015	0.030	0.630	
FIRSTLOSS	-0.101	0.051	0.047	**	-0.102	0.051	0.045	**	-0.102	0.051	0.045	**
LOSS_SEQ	-0.142	0.016	0.000	***	-0.141	0.016	0.000	***	-0.141	0.016	0.000	***
DIVDUM	0.178	0.038	0.000	***	0.168	0.038	0.000	***	0.168	0.038	0.000	***
LEVERAGE	0.275	0.087	0.002	***	0.255	0.087	0.003	***	0.255	0.087	0.003	***
T_Q	0.000	0.008	0.965		0.000	0.008	0.967		0.000	0.008	0.967	
CAPX	-0.283	0.181	0.119		-0.260	0.182	0.152		-0.260	0.182	0.152	
SPECIAL_ITEM	-5.845	0.300	0.000	***	-5.841	0.300	0.000	***	-5.841	0.300	0.000	***
CREDIT_RATING	0.193	0.034	0.000	***	0.190	0.034	0.000	***	0.190	0.034	0.000	***
DIVERSITY_D	0.071	0.041	0.083	*								
HERF_ASSET					-0.295	0.088	0.001	***				
HERF_SALES									-0.295	0.088	0.001	***
AB_D	-0.155	0.081	0.056	*	0.519	0.280	0.064	*	0.519	0.280	0.064	*
AB_D*DIVERSITY_D	0.397	0.145	0.006	***								
AB_D*HERF_ASSET					-0.637	0.311	0.041	**				
AB_D*HERF_SALES									-0.637	0.311	0.041	**
Marginal effects	0.082	0.009	0.000	***	0.132	0.007	0.000	***	0.128	0.007	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Independent variables follow Joos and Plesko (2005) with additional firm level control variables and definitions are provided in the notes to Table 4.1. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if model-predicted sales growth decreases comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 20662. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.5

Loss reversal and firm structure

Panel A: Alternative logistic regression model

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.215	0.055	0.000	***	0.005	0.080	0.953		0.042	0.081	0.607	
ACCRUAL	3.465	0.152	0.000	***	3.459	0.152	0.000	***	3.457	0.152	0.000	***
CFO	3.898	0.140	0.000	***	3.889	0.139	0.000	***	3.888	0.139	0.000	***
PAST_ACCR	-0.367	0.140	0.009	***	-0.360	0.140	0.010	**	-0.359	0.140	0.010	**
PAST_CFO	0.351	0.091	0.000	***	0.351	0.091	0.000	***	0.350	0.091	0.000	***
SIZE	0.009	0.007	0.184		0.006	0.007	0.326		0.006	0.007	0.389	
SALESGROWTH	0.090	0.024	0.000	***	0.089	0.024	0.000	***	0.089	0.024	0.000	***
FIRSTLOSS	0.016	0.036	0.646		0.016	0.036	0.653		0.016	0.036	0.655	
LOSS_SEQ	-0.140	0.012	0.000	***	-0.139	0.012	0.000	***	-0.139	0.012	0.000	***
DIVDUM	0.162	0.027	0.000	***	0.154	0.027	0.000	***	0.153	0.027	0.000	***
LEVERAGE	-0.129	0.054	0.016	**	-0.145	0.054	0.007	***	-0.148	0.053	0.006	***
T_Q	0.008	0.008	0.336		0.008	0.008	0.336		0.008	0.008	0.321	
CAPX	-0.254	0.157	0.106		-0.237	0.157	0.133		-0.237	0.157	0.132	
SPECIAL_ITEM	-5.320	0.207	0.000	***	-5.317	0.207	0.000	***	-5.319	0.207	0.000	***
CREDIT_RATING	0.178	0.024	0.000	***	0.177	0.024	0.000	***	0.176	0.024	0.000	***
DIVERSITY_D	0.046	0.028	0.092	*								
HERF_ASSET					-0.217	0.058	0.000	***				
HERF_SALES									-0.253	0.059	0.000	***
Marginal effects	0.009	0.007	0.207		-0.040	0.014	0.005	***	-0.047	0.014	0.001	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variables definitions are provided in the notes to Table 4.1. To be different from the main analysis, we decompose the ROA into CFO and accruals. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.5
Loss reversal and firm structure

Panel B: Alternative logistic regression model full model

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.104	0.058	0.073	*	0.206	0.091	0.024	**	0.248	0.092	0.007	***
ACCRUAL	3.232	0.155	0.000	***	3.227	0.155	0.000	***	3.226	0.155	0.000	***
CFO	3.819	0.142	0.000	***	3.809	0.142	0.000	***	3.809	0.142	0.000	***
PAST_ACCR	-0.321	0.146	0.028	**	-0.310	0.146	0.033	**	-0.315	0.146	0.031	**
PAST_CFO	0.416	0.094	0.000	***	0.416	0.094	0.000	***	0.414	0.094	0.000	***
SIZE	-0.004	0.007	0.601		-0.006	0.007	0.349		-0.007	0.007	0.311	
SALESGROWTH	0.047	0.025	0.061	*	0.047	0.025	0.062	*	0.047	0.025	0.061	*
FIRSTLOSS	0.000	0.037	0.998		-0.001	0.037	0.975		-0.001	0.037	0.976	
LOSS_SEQ	-0.157	0.012	0.000	***	-0.156	0.012	0.000	***	-0.156	0.012	0.000	***
DIVDUM	0.169	0.028	0.000	***	0.160	0.028	0.000	***	0.159	0.028	0.000	***
LEVERAGE	0.179	0.057	0.002	***	0.161	0.057	0.004	***	0.160	0.056	0.005	***
T_Q	0.002	0.009	0.843		0.002	0.009	0.849		0.002	0.009	0.825	
CAPX	-0.714	0.165	0.000	***	-0.696	0.165	0.000	***	-0.695	0.165	0.000	***
SPECIAL_ITEM	-6.051	0.220	0.000	***	-6.054	0.220	0.000	***	-6.055	0.220	0.000	***
CREDIT_RATING	0.202	0.025	0.000	***	0.201	0.025	0.000	***	0.200	0.025	0.000	***
DIVERSITY_D	-0.002	0.034	0.944									
HERF_ASSET					-0.095	0.073	0.192					
HERF_SALES									-0.138	0.074	0.060	*
AB_D	-1.258	0.033	0.000	***	-0.654	0.108	0.000	***	-0.691	0.110	0.000	***
AB_D*DIVERSITY_D	0.251	0.058	0.000	***								
AB_D*HERF_ASSET					-0.612	0.121	0.000	***				
AB_D*HERF_SALES									-0.567	0.122	0.000	***
Marginal effects	0.026	0.008	0.001	***	-0.072	0.014	0.000	***	-0.047	0.014	0.001	***

This table presents the results of loss reversal probability using the annual estimation of logistic regression. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. To be different from the main analysis, we decompose the ROA into CFO and accruals. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.5

Loss reversal and firm structure

Panel C: Alternative logistic regression model with high agency problem (Tobin's Q)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	-0.048	0.104	0.644		-0.091	0.141	0.518		-0.009	0.141	0.949	
ACCRUAL	3.049	0.248	0.000	***	3.032	0.248	0.000	***	3.026	0.248	0.000	***
CFO	4.026	0.241	0.000	***	4.005	0.240	0.000	***	4.000	0.240	0.000	***
PAST_ACCR	0.046	0.233	0.844		0.051	0.232	0.827		0.048	0.232	0.835	
PAST_CFO	0.811	0.171	0.000	***	0.807	0.171	0.000	***	0.804	0.171	0.000	***
SIZE	-0.020	0.010	0.050	**	-0.023	0.010	0.024	**	-0.024	0.010	0.018	**
SALESGROWTH	0.059	0.043	0.170		0.058	0.043	0.177		0.057	0.043	0.181	
FIRSTLOSS	0.054	0.052	0.296		0.052	0.052	0.315		0.052	0.052	0.311	
LOSS_SEQ	-0.148	0.018	0.000	***	-0.148	0.018	0.000	***	-0.147	0.018	0.000	***
DIVDUM	0.142	0.041	0.001	***	0.134	0.041	0.001	***	0.131	0.041	0.001	***
LEVERAGE	0.357	0.108	0.001	***	0.340	0.108	0.002	***	0.337	0.108	0.002	***
T_Q	0.233	0.073	0.001	***	0.234	0.073	0.001	***	0.236	0.073	0.001	***
CAPX	-1.311	0.336	0.000	***	-1.287	0.336	0.000	***	-1.283	0.335	0.000	***
SPECIAL_ITEM	-5.952	0.349	0.000	***	-5.949	0.349	0.000	***	-5.948	0.349	0.000	***
CREDIT_RATING	0.193	0.035	0.000	***	0.192	0.035	0.000	***	0.191	0.035	0.000	***
DIVERSITY_D	-0.072	0.049	0.145									
HERF_ASSET					0.046	0.103	0.655					
HERF_SALES									-0.042	0.103	0.683	
AB_D	-1.162	0.045	0.000	***	-0.733	0.145	0.000	***	-0.803	0.146	0.000	***
AB_D*DIVERSITY_D	0.167	0.079	0.034	**								
AB_D*HERF_ASSET					-0.439	0.162	0.007	***				
AB_D*HERF_SALES									-0.357	0.164	0.029	**
Marginal effects	0.011	0.011	0.328		-0.048	0.023	0.033	**	-0.050	0.023	0.029	**

This table presents the results of loss reversal probability with high agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. To be different from the main analysis, we decompose the ROA into CFO and accruals. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 4.5
Loss reversal and firm structure

Panel D: Logistic regression model with low agency problem (Tobin's Q)

	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
INTERCEPT	0.013	0.080	0.870		0.254	0.130	0.050	*	0.261	0.130	0.046	**
ACCRUAL	3.300	0.197	0.000	***	3.302	0.197	0.000	***	3.300	0.196	0.000	***
CFO	3.746	0.175	0.000	***	3.742	0.175	0.000	***	3.743	0.175	0.000	***
PAST_ACCR	-0.424	0.189	0.025	**	-0.411	0.190	0.030	**	-0.419	0.189	0.027	**
PAST_CFO	0.256	0.114	0.025	**	0.260	0.114	0.023	**	0.256	0.114	0.025	**
SIZE	0.002	0.009	0.822		-0.001	0.009	0.955		-0.001	0.009	0.941	
SALESGROWTH	0.011	0.032	0.722		0.012	0.032	0.715		0.012	0.032	0.706	
FIRSTLOSS	-0.070	0.053	0.188		-0.070	0.053	0.186		-0.072	0.053	0.175	
LOSS_SEQ	-0.159	0.017	0.000	***	-0.157	0.017	0.000	***	-0.158	0.017	0.000	***
DIVDUM	0.198	0.040	0.000	***	0.185	0.040	0.000	***	0.188	0.040	0.000	***
LEVERAGE	0.500	0.092	0.000	***	0.485	0.092	0.000	***	0.487	0.092	0.000	***
T_Q	-0.004	0.009	0.669		-0.004	0.009	0.672		-0.004	0.009	0.687	
CAPX	-0.641	0.194	0.001	***	-0.623	0.194	0.001	***	-0.624	0.194	0.001	***
SPECIAL_ITEM	-6.322	0.292	0.000	***	-6.320	0.292	0.000	***	-6.320	0.292	0.000	***
CREDIT_RATING	0.210	0.036	0.000	***	0.208	0.036	0.000	***	0.206	0.036	0.000	***
DIVERSITY_D	0.061	0.048	0.202									
HERF_ASSET					-0.233	0.105	0.027	**				
HERF_SALES									-0.238	0.106	0.025	**
AB_D	-1.387	0.052	0.000	***	-0.474	0.167	0.005	***	-0.468	0.170	0.006	***
AB_D*DIVERSITY_D	0.402	0.089	0.000	***								
AB_D*HERF_ASSET					-0.914	0.187	0.000	***				
AB_D*HERF_SALES									-0.915	0.190	0.000	***
Marginal effects	0.041	0.011	0.000	***	-0.094	0.018	0.000	***	-0.095	0.019	0.000	***

This table presents the results of loss reversal probability with no potential agency problems using the annual estimation of logistic regression. Over-investment is used as the proxy of agency problem which is defined as if the firm's Tobin's Q is below the Industry median in annual basis. Dependent variable is loss reversal which takes value of 1 if firms becomes profitable in the next year, and 0 otherwise. Variable definitions are provided in the notes to Table 4.1. To be different from the main analysis, we decompose the ROA into CFO and accruals. Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. AB_D (abandonment option) is dummy variable equals one if both sales and number of employees decrease comparing to the previous year, and zero otherwise. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: agency problem hypothesis under propensity matched sample

Panel A: Group by free-cash flow

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.100	0.041	0.015	**	0.102	0.041	0.014	**	0.102	0.041	0.013	**
SIZE	0.019	0.005	0.000	***	0.018	0.005	0.000	***	0.018	0.005	0.000	***
SALESGROWTH	-0.005	0.008	0.510		-0.005	0.008	0.507		-0.005	0.008	0.500	
LEVERAGE	-0.068	0.026	0.008	***	-0.070	0.026	0.006	***	-0.071	0.026	0.006	***
CAPX	-0.076	0.036	0.035		-0.075	0.036	0.035	**	-0.076	0.036	0.033	**
T_Q	-0.018	0.004	0.000	***	-0.018	0.004	0.000	***	-0.018	0.004	0.000	***
CREDIT_RATING	0.002	0.010	0.825		0.002	0.010	0.807		0.002	0.010	0.825	
CASH HOLDING	-0.049	0.027	0.063		-0.049	0.027	0.065	*	-0.049	0.026	0.064	*
RETAINED EARNINGS	0.012	0.006	0.073		0.012	0.006	0.070	*	0.012	0.007	0.070	*
R&D	0.007	0.014	0.597		0.009	0.015	0.525		0.008	0.015	0.611	
CFO	0.026	0.019	0.178		0.025	0.019	0.190		0.026	0.019	0.173	
DIVERSITY_D	0.021	0.008	0.009	***								
HERF_ASSET					-0.054	0.017	0.001	***				
HERF_SALES									-0.056	0.016	0.001	***
R^2		0.457				0.457				0.457		

This table presents the results by agency problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. Firm and year fixed effect are controlled in the regression. The number of observation is 23253. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: agency problem hypothesis under propensity matched sample

Panel B: Group by free-cash flow

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.099	0.030	0.001	***	0.098	0.030	0.001	***	0.100	0.030	0.001	***
SIZE	0.042	0.009	0.000	***	0.040	0.009	0.000	***	0.040	0.009	0.000	***
SALESGROWTH	0.004	0.009	0.663		0.003	0.009	0.705		0.004	0.009	0.682	
LEVERAGE	-0.103	0.044	0.020	**	-0.108	0.044	0.015	**	-0.107	0.044	0.016	**
CAPX	0.104	0.067	0.125		0.103	0.067	0.126		0.103	0.067	0.125	
T_Q	-0.006	0.003	0.042	**	-0.006	0.003	0.053	*	-0.006	0.003	0.049	**
CREDIT_RATING	0.013	0.018	0.467		0.013	0.018	0.485		0.013	0.018	0.488	
CASH HOLDING	-0.021	0.036	0.558		-0.020	0.036	0.574		-0.019	0.036	0.590	
RETAINED EARNINGS	-0.006	0.002	0.005	***	-0.006	0.002	0.005	***	-0.006	0.002	0.005	***
R&D	0.007	0.005	0.160		0.007	0.005	0.142		0.006	0.005	0.235	
CFO	0.033	0.023	0.158		0.033	0.023	0.152		0.033	0.023	0.150	
DIVERSITY_D	0.008	0.016	0.619									
HERF_ASSET					-0.058	0.034	0.091	*				
HERF_SALES									-0.052	0.036	0.148	
R^2		0.365				0.365				0.365		

This table presents the results by agency problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 12565. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: agency problem hypothesis under propensity matched sample

Panel C: Group by Tobin's Q

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.219	0.037	0.000	***	0.219	0.037	0.000	***	0.219	0.037	0.000	***
SIZE	0.028	0.005	0.000	***	0.027	0.005	0.000	***	0.027	0.005	0.000	***
SALESGROWTH	0.004	0.006	0.459		0.004	0.006	0.485		0.004	0.006	0.474	
LEVERAGE	-0.157	0.028	0.000	***	-0.160	0.028	0.000	***	-0.160	0.028	0.000	***
CAPX	0.029	0.041	0.483		0.030	0.041	0.460		0.030	0.041	0.462	
T_Q	-0.062	0.010	0.000	***	-0.061	0.010	0.000	***	-0.061	0.010	0.000	***
CREDIT_RATING	0.023	0.012	0.064	*	0.023	0.012	0.062	*	0.023	0.012	0.064	*
CASH HOLDING	-0.066	0.027	0.015		-0.066	0.027	0.015	**	-0.064	0.027	0.018	**
RETAINED EARNINGS	-0.001	0.003	0.774		-0.001	0.003	0.783		-0.001	0.003	0.791	
R&D	0.019	0.020	0.349		0.020	0.021	0.322		0.018	0.021	0.373	
CFO	0.020	0.021	0.350		0.019	0.021	0.358		0.020	0.021	0.351	
DIVERSITY_D	0.021	0.008	0.008	***								
HERF_ASSET					-0.057	0.016	0.000	***				
HERF_SALES									-0.055	0.017	0.002	***
R^2		0.360				0.360				0.360		

This table presents the results by agency problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 25999. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: agency problem hypothesis under propensity matched sample

Panel D: Group by Tobin's Q

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	-0.102	0.045	0.023	**	-0.102	0.045	0.023	**	-0.101	0.045	0.024	**
SIZE	-0.003	0.009	0.769		-0.003	0.009	0.734		-0.003	0.009	0.715	
SALESGROWTH	-0.007	0.016	0.662		-0.007	0.016	0.654		-0.007	0.016	0.651	
LEVERAGE	0.051	0.051	0.311		0.049	0.050	0.328		0.048	0.050	0.336	
CAPX	-0.044	0.053	0.410		-0.044	0.053	0.405		-0.045	0.053	0.400	
T_Q	-0.004	0.002	0.126		-0.004	0.002	0.132		-0.003	0.002	0.143	
CREDIT_RATING	-0.031	0.019	0.105		-0.030	0.019	0.107		-0.030	0.019	0.105	
CASH HOLDING	-0.016	0.029	0.572		-0.016	0.029	0.585		-0.016	0.029	0.578	
RETAINED EARNINGS	0.004	0.003	0.112		0.004	0.003	0.109		0.005	0.003	0.104	
R&D	0.001	0.004	0.878		0.001	0.004	0.819		0.000	0.004	0.921	
CFO	0.020	0.034	0.555		0.020	0.034	0.556		0.021	0.034	0.545	
DIVERSITY_D	0.002	0.018	0.932									
HERF_ASSET					-0.018	0.037	0.622					
HERF_SALES									-0.025	0.035	0.475	
R^2		0.523				0.523				0.523		

This table presents the results by agency problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 10400. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: financial constraint hypothesis under propensity matched sample

Panel E: Group by coverage ratio

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	-0.005	0.040	0.892		-0.005	0.040	0.905		-0.005	0.040	0.903	
SIZE	0.013	0.007	0.066	*	0.012	0.007	0.079	*	0.013	0.007	0.075	*
SALESGROWTH	-0.001	0.009	0.871		-0.002	0.009	0.865		-0.002	0.009	0.863	
LEVERAGE	-0.115	0.032	0.000	***	-0.118	0.032	0.000	***	-0.117	0.033	0.000	***
CAPX	-0.135	0.051	0.008		-0.135	0.050	0.007	***	-0.133	0.051	0.008	***
T_Q	-0.014	0.005	0.002	***	-0.014	0.005	0.003	***	-0.014	0.005	0.003	***
CREDIT_RATING	0.008	0.016	0.631		0.008	0.016	0.615		0.008	0.016	0.631	
CASH HOLDING	-0.018	0.030	0.550		-0.017	0.031	0.567		-0.017	0.030	0.567	
RETAINED EARNINGS	-0.003	0.003	0.295		-0.003	0.003	0.318		-0.003	0.003	0.303	
R&D	0.003	0.009	0.774		0.004	0.009	0.673		0.002	0.009	0.823	
CFO	0.060	0.029	0.038		0.059	0.029	0.039	**	0.060	0.029	0.036	**
DIVERSITY_D	0.022	0.013	0.096	*								
HERF_ASSET					-0.055	0.027	0.038	**				
HERF_SALES									-0.042	0.027	0.123	
R^2		0.373				0.373				0.373		

This table presents the results by financial constraint problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 25878. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: financial constraint hypothesis under propensity matched sample

Panel F: Group by coverage ratio

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.111	0.032	0.000	***	0.111	0.032	0.000	***	0.113	0.032	0.000	***
SIZE	0.026	0.005	0.000	***	0.026	0.005	0.000	***	0.025	0.005	0.000	***
SALESGROWTH	0.001	0.007	0.929		0.000	0.007	0.950		0.001	0.007	0.936	
LEVERAGE	-0.077	0.027	0.005	***	-0.079	0.027	0.004	***	-0.081	0.027	0.003	***
CAPX	0.018	0.042	0.668		0.019	0.042	0.655		0.018	0.042	0.663	
T_Q	-0.011	0.004	0.004	***	-0.010	0.004	0.004	***	-0.010	0.004	0.004	***
CREDIT_RATING	0.010	0.012	0.375		0.010	0.012	0.373		0.010	0.012	0.380	
CASH HOLDING	-0.048	0.024	0.047	**	-0.047	0.024	0.050	*	-0.047	0.024	0.051	*
RETAINED EARNINGS	-0.002	0.002	0.474		-0.002	0.002	0.469		-0.002	0.002	0.482	
R&D	0.013	0.030	0.660		0.016	0.030	0.598		0.014	0.031	0.640	
CFO	0.012	0.018	0.501		0.012	0.018	0.520		0.012	0.018	0.485	
DIVERSITY_D	0.019	0.009	0.034	**								
HERF_ASSET					-0.054	0.020	0.006	***				
HERF_SALES									-0.060	0.021	0.003	***
R^2		0.458				0.458				0.458		

This table presents the results by financial constraint problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 10376. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: financial constraint hypothesis under propensity matched sample

Panel G: Group by cash flow volatility

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.063	0.034	0.060	*	0.063	0.034	0.061	*	0.064	0.034	0.059	*
SIZE	0.029	0.010	0.003	***	0.029	0.010	0.004	***	0.029	0.010	0.003	***
SALESGROWTH	0.007	0.009	0.450		0.007	0.009	0.445		0.007	0.009	0.445	
LEVERAGE	-0.067	0.052	0.192		-0.067	0.051	0.191		-0.067	0.051	0.188	
CAPX	-0.023	0.068	0.732		-0.022	0.068	0.742		-0.022	0.068	0.745	
T_Q	-0.004	0.002	0.054	*	-0.004	0.002	0.056	*	-0.004	0.002	0.056	*
CREDIT_RATING	0.011	0.023	0.623		0.011	0.023	0.623		0.011	0.023	0.623	
CASH HOLDING	-0.042	0.044	0.345		-0.041	0.044	0.352		-0.041	0.044	0.351	
RETAINED EARNINGS	-0.005	0.002	0.032		-0.005	0.002	0.032	**	-0.005	0.002	0.036	**
R&D	0.027	0.034	0.425		0.028	0.034	0.418		0.027	0.034	0.434	
CFO	0.031	0.025	0.213		0.031	0.025	0.214		0.031	0.025	0.211	
DIVERSITY_D	0.012	0.018	0.506									
HERF_ASSET					-0.018	0.036	0.626					
HERF_SALES									-0.018	0.038	0.638	
R^2		0.392				0.392				0.392		

This table presents the results by financial constraint problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 29477. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.1

Dividend payout ratio and firm structure: financial constraint hypothesis under propensity matched sample

Panel H: Group by cash flow volatility

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.115	0.035	0.001	***	0.116	0.035	0.001	***	0.117	0.035	0.001	***
SIZE	0.026	0.005	0.000	***	0.025	0.005	0.000	***	0.025	0.005	0.000	***
SALESGROWTH	-0.002	0.008	0.835		-0.002	0.008	0.816		-0.002	0.008	0.824	
LEVERAGE	-0.080	0.026	0.002	***	-0.083	0.026	0.002	***	-0.083	0.026	0.002	***
CAPX	-0.077	0.036	0.030	**	-0.076	0.035	0.032	**	-0.077	0.035	0.030	**
T_Q	-0.021	0.005	0.000	***	-0.021	0.005	0.000	***	-0.021	0.005	0.000	***
CREDIT_RATING	0.010	0.010	0.347		0.010	0.010	0.333		0.009	0.010	0.349	
CASH HOLDING	-0.025	0.025	0.320		-0.024	0.025	0.328		-0.024	0.025	0.333	
RETAINED EARNINGS	0.001	0.004	0.836		0.001	0.004	0.850		0.001	0.004	0.854	
R&D	0.003	0.013	0.790		0.005	0.013	0.684		0.003	0.013	0.809	
CFO	0.016	0.020	0.417		0.016	0.020	0.428		0.016	0.020	0.416	
DIVERSITY_D	0.023	0.008	0.004	***								
HERF_ASSET					-0.062	0.018	0.001	***				
HERF_SALES									-0.060	0.017	0.001	***
R^2		0.343				0.343				0.343		

This table presents the results by financial constraint problem under propensity score matched sample. Propensity score matched firms are selected based on propensity scores estimated using our regression model parameters. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 7145. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: agency problem hypothesis and second stage of Heckman test

Panel A: Group by free-cash flow

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.057	0.021	0.008	***	0.057	0.021	0.007	***	0.057	0.021	0.007	***
SIZE	0.014	0.004	0.001	***	0.014	0.004	0.002	***	0.013	0.004	0.002	***
SALESGROWTH	-0.008	0.006	0.173		-0.008	0.006	0.174		-0.008	0.006	0.173	
LEVERAGE	-0.057	0.020	0.006	***	-0.059	0.020	0.004	***	-0.059	0.021	0.004	***
CAPX	-0.063	0.027	0.017	**	-0.063	0.026	0.017	**	-0.063	0.026	0.017	**
T_Q	-0.009	0.002	0.000	***	-0.009	0.002	0.000	***	-0.009	0.002	0.000	***
CREDIT_RATING	0.008	0.008	0.273		0.009	0.008	0.263		0.008	0.008	0.273	
CASH HOLDING	-0.007	0.011	0.502		-0.007	0.011	0.504		-0.007	0.011	0.492	
RETAINED EARNINGS	0.000	0.002	0.854		0.000	0.002	0.898		0.000	0.002	0.902	
R&D	0.000	0.000	0.765		0.000	0.000	0.747		0.000	0.000	0.754	
CFO	0.017	0.010	0.097	*	0.017	0.010	0.104		0.017	0.010	0.095	*
INVERSE MILLS RATIO	-0.015	0.024	0.543		-0.014	0.024	0.558		-0.015	0.024	0.531	
DIVERSITY_D	0.018	0.008	0.018	**								
HERF_ASSET					-0.048	0.017	0.004	***				
HERF_SALES									-0.048	0.016	0.002	***
R^2		0.445				0.445				0.445		

This table presents the results by agency problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 31216. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: agency problem hypothesis and second stage of Heckman test

Panel B: Group by free-cash flow

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.036	0.012	0.003	***	0.036	0.012	0.003	***	0.036	0.012	0.003	***
SIZE	0.028	0.006	0.000	***	0.027	0.006	0.000	***	0.027	0.006	0.000	***
SALESGROWTH	-0.002	0.004	0.603		-0.002	0.004	0.566		-0.002	0.004	0.577	
LEVERAGE	-0.071	0.025	0.005	**	-0.075	0.025	0.003	***	-0.075	0.025	0.003	***
CAPX	0.073	0.039	0.063		0.073	0.039	0.063	*	0.073	0.039	0.063	*
T_Q	-0.003	0.001	0.000	**	-0.003	0.001	0.000	***	-0.003	0.001	0.000	***
CREDIT_RATING	0.013	0.015	0.397		0.013	0.015	0.401		0.013	0.015	0.405	
CASH HOLDING	-0.002	0.014	0.895		-0.001	0.014	0.941		-0.001	0.014	0.939	
RETAINED EARNINGS	-0.003	0.001	0.025	***	-0.003	0.001	0.025	**	-0.003	0.001	0.024	**
R&D	0.000	0.000	0.775		0.000	0.000	0.775		0.000	0.000	0.776	
CFO	0.020	0.010	0.037		0.021	0.010	0.034	**	0.021	0.010	0.034	**
INVERSE MILLS RATIO	-0.008	0.006	0.195		-0.008	0.006	0.216		-0.008	0.006	0.215	
DIVERSITY_D	0.001	0.012	0.964									
HERF_ASSET					-0.034	0.029	0.244					
HERF_SALES									-0.032	0.031	0.292	
R^2		0.351				0.351				0.351		

This table presents the results by agency problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 19460. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.3

Dividend payout ratio and firm structure: agency problem hypothesis and second stage of Heckman test

Panel C: Group by Tobin's Q

	High Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.138	0.021	0.000	***	0.138	0.021	0.000	***	0.138	0.021	0.000	***
SIZE	0.023	0.005	0.000	***	0.022	0.005	0.000	***	0.022	0.005	0.000	***
SALESGROWTH	0.002	0.004	0.734		0.001	0.005	0.766		0.001	0.005	0.755	
LEVERAGE	-0.125	0.020	0.000	***	-0.128	0.019	0.000	***	-0.128	0.019	0.000	***
CAPX	0.002	0.029	0.933		0.002	0.029	0.934		0.003	0.029	0.928	
T_Q	-0.040	0.007	0.000	***	-0.040	0.007	0.000	***	-0.040	0.007	0.000	***
CREDIT_RATING	0.024	0.009	0.012	**	0.024	0.009	0.011	**	0.024	0.009	0.012	**
CASH HOLDING	-0.020	0.017	0.245		-0.019	0.017	0.264		-0.019	0.017	0.267	
RETAINED EARNINGS	-0.005	0.002	0.001		-0.005	0.002	0.001	***	-0.005	0.002	0.001	***
R&D	0.000	0.000	0.850		0.000	0.000	0.858		0.000	0.000	0.858	
CFO	0.022	0.014	0.132		0.021	0.014	0.134		0.022	0.014	0.131	
INVERSE MILLS RATIO	0.009	0.013	0.495		0.010	0.013	0.456		0.009	0.013	0.478	
DIVERSITY_D	0.014	0.007	0.036	**								
HERF_ASSET					-0.046	0.015	0.003	***				
HERF_SALES									-0.044	0.016	0.007	***
R^2		0.332				0.332				0.332		

This table presents the results by agency problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 34119. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: agency problem hypothesis and second stage of Heckman test

Panel D: Group by Tobin's Q

	Low Agency Problem											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	-0.017	0.010	0.098	*	-0.017	0.010	0.098	*	-0.017	0.010	0.098	*
SIZE	-0.002	0.006	0.785		-0.002	0.006	0.768		-0.002	0.006	0.751	
SALESGROWTH	-0.007	0.005	0.196		-0.007	0.005	0.193		-0.007	0.005	0.190	
LEVERAGE	0.063	0.035	0.072	*	0.063	0.035	0.074	*	0.062	0.035	0.079	*
CAPX	-0.019	0.029	0.508		-0.019	0.029	0.510		-0.019	0.029	0.508	
T_Q	0.000	0.001	0.835		0.000	0.001	0.848		0.000	0.001	0.862	
CREDIT_RATING	-0.015	0.015	0.329		-0.015	0.015	0.328		-0.015	0.015	0.326	
CASH HOLDING	-0.002	0.009	0.809		-0.002	0.009	0.820		-0.002	0.009	0.823	
RETAINED EARNINGS	0.002	0.001	0.071	*	0.002	0.001	0.071	*	0.002	0.001	0.071	*
R&D	0.000	0.000	0.588		0.000	0.000	0.588		0.000	0.000	0.587	
CFO	0.012	0.010	0.232		0.012	0.010	0.232		0.012	0.010	0.224	
INVERSE MILLS RATIO	-0.012	0.009	0.167		-0.012	0.009	0.166		-0.012	0.009	0.167	
DIVERSITY_D	0.009	0.014	0.500									
HERF_ASSET					-0.024	0.031	0.433					
HERF_SALES									-0.029	0.030	0.326	
R^2		0.492				0.492				0.492		

This table presents the results by agency problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of high agency and low agency problems. The criteria for agency problem are free cash flow (Panel A &B) and Tobin's Q (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 17031. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: financial constraint hypothesis with second stage of Heckman test

Panel E: Group by Coverage

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	-0.021	0.016	0.182		-0.021	0.016	0.186		-0.021	0.016	0.180	
SIZE	0.007	0.005	0.204		0.007	0.005	0.226		0.007	0.005	0.206	
SALESGROWTH	-0.003	0.004	0.412		-0.003	0.004	0.414		-0.003	0.004	0.421	
LEVERAGE	-0.097	0.023	0.000	***	-0.099	0.022	0.000	***	-0.097	0.023	0.000	***
CAPX	-0.072	0.034	0.036		-0.071	0.034	0.037	**	-0.071	0.034	0.039	**
T_Q	-0.006	0.002	0.001	***	-0.006	0.002	0.002	***	-0.006	0.002	0.002	***
CREDIT_RATING	0.005	0.011	0.685		0.005	0.011	0.670		0.005	0.011	0.684	
CASH HOLDING	0.002	0.010	0.827		0.002	0.010	0.818		0.002	0.010	0.852	
RETAINED EARNINGS	-0.002	0.002	0.160		-0.002	0.002	0.167		-0.002	0.002	0.168	
R&D	0.000	0.000	0.022		0.000	0.000	0.023	**	0.000	0.000	0.022	**
CFO	0.031	0.011	0.004		0.031	0.011	0.004	***	0.031	0.011	0.004	***
INVERSE MILLS RATIO	-0.020	0.017	0.253		-0.020	0.017	0.249		-0.020	0.017	0.237	
DIVERSITY_D	0.014	0.011	0.190									
HERF_ASSET					-0.038	0.022	0.085	*				
HERF_SALES									-0.025	0.023	0.282	
R^2		0.336				0.336				0.336		

This table presents the results by financial constraint problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 35678. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: financial constraint hypothesis with second stage of Heckman test

Panel F: Group by Coverage

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.053	0.012	0.000	***	0.054	0.012	0.000	***	0.054	0.012	0.000	***
SIZE	0.020	0.004	0.000	***	0.019	0.004	0.000	***	0.019	0.004	0.000	***
SALESGROWTH	-0.002	0.004	0.575		-0.002	0.004	0.571		-0.002	0.004	0.578	
LEVERAGE	-0.055	0.020	0.006	***	-0.057	0.020	0.005	***	-0.058	0.020	0.004	***
CAPX	-0.002	0.023	0.933		-0.002	0.023	0.942		-0.002	0.023	0.938	
T_Q	-0.004	0.001	0.000	***	-0.004	0.001	0.000	***	-0.004	0.001	0.000	***
CREDIT_RATING	0.012	0.010	0.197		0.013	0.010	0.191		0.012	0.010	0.195	
CASH HOLDING	-0.008	0.013	0.553		-0.008	0.014	0.565		-0.007	0.013	0.581	
RETAINED EARNINGS	-0.001	0.001	0.275		-0.001	0.001	0.288		-0.001	0.001	0.286	
R&D	0.000	0.000	0.003	***	0.000	0.000	0.003	***	0.000	0.000	0.003	***
CFO	0.010	0.009	0.239		0.010	0.009	0.250		0.011	0.009	0.227	
INVERSE MILLS RATIO	0.003	0.005	0.629		0.003	0.005	0.608		0.003	0.005	0.610	
DIVERSITY_D	0.019	0.009	0.029	**								
HERF_ASSET					-0.049	0.020	0.014	**				
HERF_SALES									-0.055	0.020	0.006	***
R^2		0.454				0.454				0.454		

This table presents the results by financial constraint problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 15463. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: financial constraint hypothesis with second stage of Heckman test

Panel G: Group by cash flow volatility

	Unconstraint Firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.031	0.013	0.016	**	0.031	0.013	0.015	**	0.031	0.013	0.016	**
SIZE	0.015	0.006	0.016	**	0.015	0.006	0.018	**	0.015	0.006	0.016	**
SALESGROWTH	0.000	0.004	0.954		0.000	0.004	0.951		0.000	0.004	0.950	
LEVERAGE	-0.051	0.030	0.093	*	-0.051	0.030	0.090	*	-0.051	0.030	0.085	*
CAPX	0.005	0.037	0.897		0.005	0.037	0.896		0.005	0.037	0.895	
T_Q	-0.002	0.001	0.053	*	-0.002	0.001	0.053	*	-0.002	0.001	0.054	*
CREDIT_RATING	0.014	0.013	0.284		0.014	0.013	0.283		0.014	0.013	0.283	
CASH HOLDING	-0.020	0.011	0.074		-0.020	0.011	0.073	*	-0.020	0.011	0.075	*
RETAINED EARNINGS	-0.001	0.002	0.433		-0.001	0.002	0.431		-0.001	0.002	0.431	
R&D	0.000	0.000	0.490		0.000	0.000	0.490		0.000	0.000	0.491	
CFO	0.013	0.010	0.204		0.013	0.010	0.203		0.013	0.010	0.203	
INVERSE MILLS RATIO	-0.006	0.005	0.288		-0.006	0.005	0.287		-0.006	0.005	0.287	
DIVERSITY_D	0.004	0.015	0.791									
HERF_ASSET					-0.010	0.031	0.749					
HERF_SALES									-0.009	0.034	0.795	
R^2		0.376				0.376				0.376		

This table presents the results by financial constraint problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 37811. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 5.2

Dividend payout ratio and firm structure: financial constraint hypothesis with second stage of Heckman test

Panel H: Group by cash flow volatility

	Constraint firms											
	(1)				(2)				(3)			
	Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value		Coef.	Std. Err.	p-value	
ROA	0.047	0.014	0.001	***	0.048	0.014	0.001	***	0.048	0.014	0.001	***
SIZE	0.020	0.004	0.000	***	0.020	0.004	0.000	***	0.019	0.004	0.000	***
SALESGROWTH	-0.005	0.005	0.356		-0.005	0.005	0.350		-0.005	0.005	0.357	
LEVERAGE	-0.066	0.019	0.001	***	-0.069	0.019	0.000	***	-0.069	0.019	0.000	***
CAPX	-0.077	0.029	0.009	***	-0.076	0.029	0.009	***	-0.076	0.029	0.009	***
T_Q	-0.008	0.002	0.000	***	-0.008	0.002	0.000	***	-0.008	0.002	0.000	***
CREDIT_RATING	0.009	0.009	0.343		0.009	0.009	0.325		0.009	0.009	0.344	
CASH HOLDING	0.009	0.012	0.443		0.009	0.012	0.430		0.009	0.012	0.437	
RETAINED EARNINGS	-0.002	0.001	0.138		-0.002	0.001	0.149		-0.002	0.001	0.144	
R&D	0.000	0.000	0.185		0.000	0.000	0.190		0.000	0.000	0.190	
CFO	0.018	0.011	0.110		0.018	0.011	0.112		0.018	0.011	0.108	
INVERSE MILLS RATIO	0.002	0.020	0.937		0.003	0.020	0.895		0.002	0.020	0.937	
DIVERSITY_D	0.019	0.007	0.009	***								
HERF_ASSET					-0.055	0.017	0.001	**				
HERF_SALES									-0.053	0.017	0.002	***
R^2		0.285				0.285				0.285		

This table presents the results by financial constraint problem under the second stage of Heckman test. We obtain Inverse Mills Ratio from the first stage of Heckman test using a probit model in Table 8 Panel A. The regressions are presented across the groups of financially unconstraint and constraint. The criteria for financial constraint are coverage ratio (Panel A &B) and cash flow volatility (Panel C &D). Three measures of diversification are used. In regression (1), diversification is dummy variable equals one if a firm has more than one segments and zero otherwise. In regression (2) and (3), diversification is calculated as Herfindahl Index based on asset and sales. Dependent variable is the dividend payout ratio defined as total dividend divided by net income. Definitions of independent variables are provided in the notes to Table 5.1. Reported coefficients are the average coefficients over the estimation period from 1980 to 2016. The number of observation is 13552. Firm and year fixed effect are controlled in the regression. Standard errors are clustered by firm and fiscal year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.